

STIC-Biotech/ChemLib

83852

Fr m: Snedden, Sheridan
Sent: Wednesday, January 08, 2003 2:48 PM
T : STIC-Biotech/ChemLib
Subject: Oligo Search Request 09424815

Sheridan SNEDDEN ID# 79298 Date: 1/ 7 /2002

AU 1653

308-4843

Serial # : 09424815

Room Location: 10A12

Mail Box: 9B01

Earliest Priority Filing Date: 05/18/1998

Please perform an **OLIGO** search for SEQ ID NO: 1

Thanks,
Examiner Snedden
#79298
A.U. 1653/ 9B01
Office Location: 10A12
Phone #: 305-4843

CRFE
Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 1/9/03
Date Completed: 1/9/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 02
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:25:14 ; Search time 35 Seconds
(without alignments)
224.622 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59
Sequence: 1 KKHGSLARAGKVRGQTPKVA.....RRFVNVVPTFGKKKGFNANS 59

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: /SIDSeq_101002: *
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	59	100.0	59	15	AAK48204	Heparin-binding pr
2	59	100.0	59	20	AAW95381	Antimicrobial pept
3	59	100.0	66	23	ABP42408	Human ovarian anti
4	59	100.0	170	21	AA843549	Human cancer assoc
5	41	69.5	41	17	AAW03681	Long term potentia
6	39	66.1	108	22	ABG37228	Novel human diagno
7	39	66.1	108	22	ABG27519	Novel human diagno
8	37	62.7	40	22	AAO08731	Human polypeptide
9	37	62.7	157	22	AAW5965	Human colon cancer
10	32	54.2	130	21	AAW01569	Human secreted pro

11	32	54.2	312	22	AAU30872	Novel human secret
12	31	52.5	61	22	ABG27230	Novel human diagno
13	31	52.5	90	22	ABG27520	Novel human diagno
14	28	47.5	118	22	ABB29449	Peptide #2100 enco
15	28	47.5	118	22	ABB34625	Peptide #2131 enco
16	28	47.5	118	22	ABB20035	Protein #2034 enco
17	28	47.5	118	22	AAW54415	Human brain expres
18	28	47.5	118	22	AAW67811	Human bone marrow
19	28	47.5	118	22	AAW15620	Peptide #2054 enco
20	28	47.5	118	22	AAW28118	Peptide #2046 enco
21	28	47.5	118	22	AAW03364	Human peptide enco
22	28	47.5	118	22	ABG37354	Human polypeptide
23	24	40.7	35	22	AAW07113	Arabidopsis thalia
24	22	37.3	62	21	AAW04311	Arabidopsis thalia
25	22	37.3	62	21	AAW15348	Arabidopsis thalia
26	22	37.3	62	21	AAW22753	Arabidopsis thalia
27	22	37.3	62	21	AAW38888	Arabidopsis thalia
28	22	37.3	62	21	AAW34119	Arabidopsis thalia
29	22	37.3	62	21	AAW38861	Arabidopsis thalia
30	22	37.3	62	21	AAW39476	Arabidopsis thalia
31	22	37.3	62	21	AAW40704	Arabidopsis thalia
32	22	37.3	62	21	AAW40788	Arabidopsis thalia
33	22	37.3	62	21	AAW44883	Arabidopsis thalia
34	22	37.3	62	21	AAW48690	Arabidopsis thalia
35	22	37.3	68	21	AAW31871	Arabidopsis thalia
36	22	37.3	82	21	AAW40703	Arabidopsis thalia
37	22	37.3	83	21	AAW44882	Arabidopsis thalia
38	22	37.3	84	21	AAW32887	Arabidopsis thalia
39	22	37.3	93	21	AAW34118	Arabidopsis thalia
40	22	37.3	110	21	AAW32886	Arabidopsis thalia
41	22	37.3	130	21	AAW26728	Arabidopsis thalia
42	21	35.6	230	23	ABG34133	Human seb4d protel
43	21	35.6	459	23	ABG34132	Human seb4d protel
44	20	33.9	65	21	AAW20231	Human seb4d protel
45	19	32.2	229	22	AAW31057	Novel human secret

ALIGNMENTS

RESULT 1
AAW48204 standard; protein: 59 AA.
ID AAK48204
AC AAK48204;
XX
DT 12-JUL-1994 (first entry)
XX
DE Heparin-binding protein.
XX
KW Cell growth agent; wound; bone disease; treatment; treating; agent.
XX
OS Rattus norvegicus.
XX
PM JP05339287-A.
XX
PD 21-DEC-1993.
XX
PF 05-JUN-1992; 92JP-0145125.
XX
PR 05-JUN-1992; 92JP-0145125.
XX
PA (FARH) HOECHST JAPAN LTD.
XX
XX WPI; 1994-031824/04.
XX
DR New heparin-binding protein - used as a cell growth agent for
XX treatment of wounds and bone disease
XX
PT Claim 1; Page 3; 4pp; Japanese.
XX
PS The sequence is that of a heparin binding-protein which is useful as
CC a cell growth agent and in the treatment of wounds and bone disease.

```

XX      SQ      Sequence      59 AA:
Query Match      100.0%; Score 59; DB 15; Length 59;
Best Local Similarity 100.0%; Pred. NO. 1.9e-54;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KVHGSILARAGKVGQTPKVAKEKKKKTKGRARRMQYNRRFVNVVPTFGKKKGNPANS 59
Db      1 KVHGSILARAGKVGQTPKVAKEKKKKTKGRARRMQYNRRFVNVVPTFGKKKGNPANS 59

RESULT 2
AAM95381
ID      AAM95381 standard; peptide; 59 AA.
XX
XX      AC      AAM95381;
XX
XX      DT      17-MAR-1999 (first entry)
XX
XX      DE      Antimicrobial peptide fragment from ubiquicidine.
XX
XX      KW      Ubiquicidine; [treatment; diagnosis; prophylaxis; infection; microbial;
KW      pathogenic; Gram-positive bacteria; antimicrobial; Staphylococcus aureus;
KW      Listeria monocytogenes; Gram-negative; Klebsiella pneumoniae; E. coli;
KW      enterococcus; Salmonella typhimurium; Mycobacterium avium; M. fortuitum;
KW      fungus; Candida albicans; Cryptococcus neoformans; Aspergillus fumigatus;
KW      virus; parasite; Trypanosoma cruzi; Toxoplasma gondii.
XX
XX      OS      Synthetic.
XX      OS      Mus sp.
XX
XX      PN      WO9854314-A1.
XX
XX      PD      03-DEC-1998.
XX
XX      PF      29-MAY-1998; 98MO-NL00311.
XX
XX      PR      29-MAY-1997; 97NL-1006164.
XX
XX      PA      (UYLE-) RIKKSUNIV LEIDEN.
XX
XX      PI      Feltzma RIJ, Hienstra PS, Nibbering PH, Pauwels EKI;
XX      PI      Van Den Barselaar MT;
XX
XX      DR      WPI: 1999-070214/06.
XX
XX      PT      New antimicrobial peptides derived from ubiquicidine - useful for
XX      PT      the prophylaxis, diagnosis and treatment of infections in humans and
XX      PT      animals
XX
XX      PS      Claim 2; Page 23; 48pp; English.
XX
XX      CC      Sequences AAM95381-389 represent antimicrobial peptide fragments derived
XX      CC      from ubiquicidine. Ubiquicidine or optionally modified peptide fragments
XX      CC      of ubiquicidine, may be used for the treatment, diagnosis, or
XX      CC      prophylaxis of infections in humans and animals. In particular the
XX      CC      products and methods are directed against microbial infections caused by
XX      CC      pathogenic Gram-positive Staphylococcus aureus, including antibiotic
XX      CC      resistant strains, Listeria monocytogenes, and Gram-negative antibiotic
XX      CC      resistant Klebsiella pneumoniae, E. coli, enterococci, and Salmonella
XX      CC      typhimurium bacteria, micro-organisms difficult to treat such as
XX      CC      Mycobacterium avium and M. fortuitum, fungi such as Candida albicans,
XX      CC      Cryptococcus neoformans, and Aspergillus fumigatus, viruses, in
XX      CC      particular enveloped viruses, and parasites such as Trypanosoma cruzi and
XX      CC      Toxoplasma gondii.
XX
XX      SQ      Sequence 59 AA:
Query Match      100.0%; Score 59; DB 20; Length 59;
Best Local Similarity 100.0%; Pred. NO. 1.9e-54;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 KVHGSILARAGKVGQTPKVAKEKKKKTKGRARRMQYNRRFVNVVPTFGKKKGNPANS 59
Db      1 KVHGSILARAGKVGQTPKVAKEKKKKTKGRARRMQYNRRFVNVVPTFGKKKGNPANS 59

RESULT 3
ABP42408
ID      ABP42408 standard; Protein; 66 AA.
XX
XX      AC      ABP42408;
XX
XX      DT      22-AUG-2002 (first entry)
XX
XX      DE      Human ovarian antigen HOCM79, SEQ ID NO:3540.
XX
XX      KW      Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW      ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW      infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW      PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW      inflammatory condition; immune disorder; blood disorder;
KW      cardiovascular disorder; respiratory disorder; neurological disorder;
KW      gastrointestinal disorder; urinary system disorder; drug screening;
KW      gene therapy; chromosome mapping; forensic analysis;
KW      antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW      antiinflammatory; gynaecological; reproductive.
XX
XX      OS      Homo sapiens.
XX
XX      PN      WO200200677-A1.
XX
XX      PD      03-JAN-2002.
XX
XX      PF      07-JUN-2001; 2001WO-US18569.
XX
XX      PR      07-JUN-2000; 2000US-209467P.
XX
XX      PA      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      PI      Blise CE, Rosen CA;
XX      PI      WPI: 2002-147878/19.
XX      PI      N-PSDB; ABQ53485.
XX
XX      DR      Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX      DR      useful in the prevention, treatment and diagnosis of cancer (e.g.
XX      DR      ovarian cancer), immune disorders, cardiovascular disorders and
XX      DR      neurological diseases -
XX
XX      PS      Claim 11; SEQ ID NO 3540; 2922pp; English.
XX
XX      CC      The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX      CC      ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX      CC      encompasses polypeptides 908 identical and polynucleotides 958 identical
XX      CC      to the sequences of the invention. The invention additionally relates to
XX      CC      recombinant vectors and host cells comprising human ovarian antigen
XX      CC      polynucleotides, antibodies against human ovarian antigens, and the use
XX      CC      of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX      CC      treating, prognosing or preventing various ovary and/or breast-related
XX      CC      disorders. Such conditions include ovarian cancer and breast cancer, and
XX      CC      metastatic tumours of ovarian or breast origin, reproductive system
XX      CC      disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX      CC      polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX      CC      disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX      CC      shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX      CC      vaginitis), immune disorders (e.g., congenital and acquired
XX      CC      immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX      CC      blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX      CC      respiratory disorders, neurological disorders, gastrointestinal disorders
XX      CC      and urinary system disorders. Ovarian antigen polypeptides and
XX      CC      polynucleotides may also be used in screening for compounds which
XX      CC      modulate ovarian antigen expression or activity. The polynucleotides may
XX      CC      further be used for gene therapy, chromosome mapping, in the
XX      CC      identification of individuals and in forensic analysis, and the

```

CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/publised_pcl_sequences.

XX Sequence 66 AA:

SO Query Match 100.0%; Score 59; DB 23; Length 66;
Best Local Similarity 100.0%; Pred. No. 2.1e-54;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRRVNVVPTFGKKKGPANNS 59
DB 8 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRRVNVVPTFGKKKGPANNS 66

RESULT 4
AAB43549

ID AAB43549 standard; Protein: 170 AA.

XX AAB43549;

XX 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:994.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
KM diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KM antiinflammatory; antihistaminic; antirheumatic; antihistolic; antiviral;
KM dermatological; neutroprotective; thrombolytic; coagulant; cardiant;
KM vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KM immune disorder; haematopoietic cell disorder; autoimmune disorder;
KM allergic reaction; graft versus host disease; organ rejection;
KM haemostatic; thrombolytic; cardiovascular disorder; infection;
KM neurological disease; drug screening.

XX Homo sapiens.

XX WO200055350-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

DR N-PSDB; AAC77758.

PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer -

PS Claim 11; Page 1568-1569; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given
CC in AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antiinflammatory; antihistaminic; antirheumatic; antihistolic;
CC antiinflammatory; antihistaminic; antirheumatic; antihistolic;
CC dermatological; neutroprotective; cardiant; thrombolytic; coagulant;
CC nocotropic; vasotrophic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating

CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.

XX Sequence 170 AA:

SO Query Match 100.0%; Score 59; DB 21; Length 170;
Best Local Similarity 100.0%; Pred. NO. 4.7e-54;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRRVNVVPTFGKKKGPANNS 59
DB 112 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRRVNVVPTFGKKKGPANNS 170

RESULT 5
AAM03681

ID AAM03681 standard; peptide: 41 AA.

XX AAM03681;

XX 31-DEC-1996 (first entry)

DE Long term potentiation inducing peptide.

XX Long term potentiation inducing peptide; mouse; brain; senile dementia;
KM cerebrovascular; Alzheimer's disease.

XX Mus musculus.

XX JP08176193-A.

XX 09-JUL-1996.

XX 23-DEC-1994; 94JP-0336143.

XX 23-DEC-1994; 94JP-0336143.

XX (SOSE-) SOSEI KK.

XX WPI; 1996-368226/37.

XX A long-term potentiation inducing peptide - useful for the diagnosis
XX and treatment of senile dementia

XX Claim 1; Page 6; 7pp; Japanese.

XX This is the sequence of a long term potentiation inducing peptide
XX isolated from mice brains by centrifugation, extraction, dialysis,
XX reverse phase HPLC and gel filtration chromatography (Sephadex C-25).

XX The peptide is useful for diagnosis and treatment of senile dementia
XX e.g. cerebrovascular or Alzheimer's dementias.

XX Sequence 41 AA:

SO Query Match 69.5%; Score 41; DB 17; Length 41;
Best Local Similarity 100.0%; Pred. NO. 1.1e-35;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRR 41
DB 1 KKHGSLARAGKVRGOTPKVAKOEKKKKTKGRARRRQYRR 41

RESULT 6
ABG27228
ID ABG27228 standard; Protein: 108 AA.

XX AC ABG27228;
 XX 18-FEB-2002 (first entry)
 DT XX
 DE Novel human diagnostic protein #27219.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS91415.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 PS Claim 20; SEQ ID No 57587; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 108 AA:
 Query Match 66.1%; Score 39; DB 22; Length 108;
 Best Local Similarity 100.0%; Pred. No. 3.2e-33;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 20 AKOEKKKKTKGAKRRMOYNRFFVNVPTFGKKKGNAN 58
 DB 69 AKOEKKKKTKGAKRRMOYNRFFVNVPTFGKKKGNAN 107

XX DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27510.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB: AAS91706.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 PS Claim 20; SEQ ID No 57878; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 108 AA:
 Query Match 66.1%; Score 39; DB 22; Length 108;
 Best Local Similarity 100.0%; Pred. No. 3.2e-33;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 20 AKOEKKKKTKGAKRRMOYNRFFVNVPTFGKKKGNAN 58
 DB 69 AKOEKKKKTKGAKRRMOYNRFFVNVPTFGKKKGNAN 107

RESULT 7
 ID ABG27519 standard; Protein; 108 AA.
 AC ABG27519;

RESULT 8
 ID AA008731 standard; Protein; 40 AA.
 AC AA008731;
 DT 06-NOV-2001 (first entry)

PN	WO200122920-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US26524.
XX	
PR	29-SEP-1999; 99US-0157137.
PR	03-NOV-1999; 99US-0163280.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Barash SC, Birse CE, Rosen CA:
XX	
DR	WP1: 2001-235357/24.
DR	N-PSDB; AAH35370.
XX	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX	
PS	useful for preventing, diagnosing and/or treating colorectal cancers -
XX	
PS	Claim 11: Page 8192-8193; 9803pp; English.
XX	
CC	AAH32943 to AAH37195 and AAG7514 to AAG77788 represent human colon
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where
CC	the proteins are collectively known as colon cancer antigens. The colon
CC	cancer antigens have cytosolic activity and can be used in gene
CC	therapy and vaccine production. N and P may be used in the prevention,
CC	diagnosis and treatment of diseases associated with inappropriate P
CC	expression. For example, N and P may be used to treat disorders
CC	associated with decreased expression by rectifying mutations or deletions
CC	in a patient's genome that affect the activity of P by expressing
CC	inactive proteins or to supplement the patients own production of P.
CC	Additionally, N may be used to produce the colon cancer-associated Ps,
CC	by inserting the nucleic acids into a host cell and culturing the cell
CC	to express the proteins. N and P can be used in the prevention, diagnosis
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC	and AAH7789 represent sequences used in the exemplification of the
CC	present invention.
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC	missing at time of publication, meaning no sequences are present for
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.
XX	
SO	Sequence 157 AA:
XX	
Query Match	62.7%; Score 37; DB 22; Length 157;
Best Local Similarity	100.0%; Pred. No. 5.7e-31;
Matches	37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 KVNGLARAGKRGOTPRVAAKOEKKKKTKGAKRRMQ 37
DB	106 KVNGLARAGKRGOTPRVAAKOEKKKKTKGAKRRMQ 142
XX	
RESULT 10	
ID	AAAG01569
XX	
AC	AAAG01569 standard; Protein: 130 AA.
XX	
AC	AAAG01569;
XX	
DT	06-OCT-2000 (first entry)
XX	
DE	Human secreted protein, SEQ ID NO: 5650.
XX	
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX	gene therapy; chromosome mapping.
OS	Homo sapiens.
XX	
XX	EP1033401-A2.
PN	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-0200610.
XX	

PR 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSEP.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI: 2000-500381/45.
DR N-PSDB: AAC01575.
XX
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
XX Claim 13; SEQ ID 5650; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 130 AA:
SQ
Query Match 54.2%; Score 32; DB 21; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.6e-26;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KYHGSILARAGKVRGQTPKVAQEKKKKTGRA 32
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 75 KYHGSILARAGKVRGQTPKVAQEKKKKTGRA 106
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
RESULT 11
AAU30872
ID AAU30872 standard; Protein; 312 AA.
XX
XX AAU30872;
AC
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Novel human secreted protein #1363.
DE
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200179449-A2.
PN
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001WO-US08656.
PF
XX
XX 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI: 2001-611725/70.
DR
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -

XX
XX Claim 20; Page 364; 765pp; English.
PS
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AAU29510-AAU3304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.
SQ Sequence 312 AA:
Query Match 54.2%; Score 32; DB 22; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.9e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 23 EKKKKKTGRKRRMOYNRRFVNVPTEGKKKG 54
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
DB 72 EKKKKKTGRKRRMOYNRRFVNVPTEGKKKG 103
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
RESULT 12
ABG27230
ID ABG27230 standard; Protein; 61 AA.
XX
XX ABG27230;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #27221.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI: 2001-639362/73.
DR N-PSDB: AAS91417.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
XX Claim 20; SEQ ID No 57589; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence.
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 118 AA:

Query Match 47.5%; Score 28; DB 22; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHGSLARAGKVRGQTPPKVAKOEKKKK 28
DB 56 KKHGSLARAGKVRGQTPPKVAKOEKKKK 83

RESULT 15

ABB34625
ID ABB34625 standard; Peptide; 118 AA.

AC ABB34625;

DT 04-FEB-2002 (first entry)

DE Peptide #2131 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -

PS Claim 27; SEQ ID NO 27260; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 118 AA:

Query Match 47.5%; Score 28; DB 22; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHGSLARAGKVRGQTPPKVAKOEKKKK 28
DB 56 KKHGSLARAGKVRGQTPPKVAKOEKKKK 83

Search completed: January 9, 2003, 12:29:33
Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:28:54 : Search time 14 Seconds
(without alignments)
123.997 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59
Sequence: 1 KVGSLRAGKVRGQTPKVA.....RRFVNVPTFGKKKKPNANS 59

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/laa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/laa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/laa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/laa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/laa/PCrUS.COMB.pep:*
6: /cgn2_6/ptodata/1/laa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	11.9	381	1 US-08-487-810-2	Sequence 2, Appl1
2	7	11.9	409	2 US-08-743-130A-2	Sequence 2, Appl1
3	7	11.9	409	2 US-08-743-130A-39	Sequence 39, Appl1
4	7	11.9	7257	3 US-09-335-409-5	Sequence 5, Appl1
5	7	11.9	7257	4 US-09-568-102-5	Sequence 5, Appl1
6	7	11.9	7257	4 US-09-567-869-5	Sequence 5, Appl1
7	7	11.9	7257	4 US-09-568-480-5	Sequence 5, Appl1
8	7	11.9	7257	4 US-09-568-486-5	Sequence 5, Appl1
9	7	11.9	7257	4 US-09-568-472-5	Sequence 5, Appl1
10	7	11.9	7257	4 US-09-567-899-5	Sequence 5, Appl1
11	6	10.2	15	4 US-09-082-358B-24	Sequence 24, Appl1
12	6	10.2	15	4 US-09-082-358B-62	Sequence 62, Appl1
13	6	10.2	16	3 US-08-851-843A-201	Sequence 201, App
14	6	10.2	16	4 US-08-974-549A-320	Sequence 320, App
15	6	10.2	16	4 US-08-854-050-201	Sequence 201, App
16	6	10.2	16	4 US-09-430-323-201	Sequence 201, App
17	6	10.2	29	4 US-09-177-349-189	Sequence 189, App
18	6	10.2	51	4 US-09-227-357-479	Sequence 479, App
19	6	10.2	55	4 US-08-569-749-5	Sequence 5, Appl1
20	6	10.2	55	4 US-08-569-749-6	Sequence 6, Appl1
21	6	10.2	55	5 PCT-US96-12860-5	Sequence 5, Appl1
22	6	10.2	55	5 PCT-US96-12860-6	Sequence 6, Appl1
23	6	10.2	57	4 US-09-082-358B-43	Sequence 43, Appl1
24	6	10.2	63	4 US-08-971-089-2	Sequence 2, Appl1
25	6	10.2	68	2 US-08-511-485-18	Sequence 18, Appl1
26	6	10.2	68	2 US-08-511-485-19	Sequence 19, Appl1
27	6	10.2	141	4 US-09-091-725-51	Sequence 51, Appl1

28	6	10.2	154	4 US-08-871-732A-1	Sequence 1, Appl1
29	6	10.2	154	4 US-09-346-510B-1	Sequence 1, Appl1
30	6	10.2	175	2 US-08-533-669A-4	Sequence 4, Appl1
31	6	10.2	175	4 US-09-183-861-4	Sequence 4, Appl1
32	6	10.2	175	4 US-09-022-765-4	Sequence 4, Appl1
33	6	10.2	197	1 US-08-468-347-24	Sequence 24, Appl1
34	6	10.2	197	2 US-08-467-389-24	Sequence 24, Appl1
35	6	10.2	197	2 US-08-779-379-24	Sequence 24, Appl1
36	6	10.2	197	2 US-08-469-219-24	Sequence 24, Appl1
37	6	10.2	197	2 US-08-228-152-24	Sequence 24, Appl1
38	6	10.2	234	2 US-08-903-801-3	Sequence 3, Appl1
39	6	10.2	234	4 US-09-295-055-3	Sequence 3, Appl1
40	6	10.2	251	4 US-09-149-476-568	Sequence 568, App
41	6	10.2	315	4 US-09-347-803-20	Sequence 20, Appl1
42	6	10.2	318	4 US-09-615-192A-389	Sequence 389, App
43	6	10.2	341	4 US-09-347-803-16	Sequence 16, Appl1
44	6	10.2	370	2 US-08-729-214-27	Sequence 27, Appl1
45	6	10.2	370	3 US-09-028-934-27	Sequence 27, Appl1

ALIGNMENTS

```
RESULT 1
US-08-487-810-2
: Sequence 2, Application US/08487810
: Patent No. 5618695
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: DNA ENCODING HEW-1, A GENE EXPRESSED BY
: TITLE OF INVENTION: SCLEROSING HEMANGIOMA CELLS
: NUMBER OF SEQUENCES: 2
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/487,810
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 381 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-487-810-2

Query Match      11.9% Score 7: DB 1: Length 381
Best Local Similarity 100.0% Pred. No. 8.5:
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY      22 QEKKKK 28
      |||||
Db      375 QEKKKK 381

RESULT 2
US-08-743-130A-2
: Sequence 2, Application US/08743130A
: Patent No. 5871987
: GENERAL INFORMATION:
: APPLICANT: Sassanfar, Mandana
: APPLICANT: Gallant, Paul L.
: APPLICANT: Shen, Xiaoyu
: APPLICANT: Tao, Mianjun
: APPLICANT: Tao, Jianshi
: APPLICANT: Houman, Fariba
: TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESS:
: ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
```

STREET: Two Millia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI95-12
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-743-130A-2

Query Match 11.9%; Score 7; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 KOEKKK 27
|||||
DB 360 KOEKKK 366

RESULT 3
US-08-743-130A-39
Sequence 39, Application US/08/43130A
Patent No. 5871987
GENERAL INFORMATION:
APPLICANT: Sassanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Tao, Jianjun
APPLICANT: Houman, Fariba
TITLE OF INVENTION: CANDIDA TYROSYL-ERNA SYNTHETASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
STREET: Two Millia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI95-12

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
US-08-743-130A-39

Query Match 11.9%; Score 7; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 KOEKKK 27
|||||
DB 360 KOEKKK 366

RESULT 4
US-09-335-409-5
Sequence 5, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-5

Query Match 11.9%; Score 7; DB 3; Length 7257;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SLARAK 11
|||||
DB 7250 SLARAK 7256

RESULT 5
US-09-568-102-5
Sequence 5, Application US/09568102
Patent No. 6346404
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 5
LENGTH: 7257

```

: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-568-102-5

```

```

Query Match      11.9%: Score 7; DB 4; Length 7257;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 SLARAK 11
      |||||||
Db      7250 SLARAK 7256

```

```

RESULT 6
US-09-567-969-5
: Sequence 5, Application US/09567969
: Patent No. 6355457
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/567,969
: CURRENT FILING DATE: 2000-05-10
: PRIOR APPLICATION NUMBER: 09/335,409
: PRIOR FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 5
: LENGTH: 7257
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-567-969-5

```

```

Query Match      11.9%: Score 7; DB 4; Length 7257;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 SLARAK 11
      |||||||
Db      7250 SLARAK 7256

```

```

RESULT 7
US-09-568-480-5
: Sequence 5, Application US/09568480
: Patent No. 6355458
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/568,480
: CURRENT FILING DATE: 2000-05-10
: PRIOR APPLICATION NUMBER: 09/335,409
: PRIOR FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 5
: LENGTH: 7257
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-568-480-5

```

```

Query Match      11.9%: Score 7; DB 4; Length 7257;

```

```

Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 SLARAK 11
      |||||||
Db      7250 SLARAK 7256

```

```

RESULT 8
US-09-568-486-5
: Sequence 5, Application US/09568486
: Patent No. 6355459
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/568,486
: CURRENT FILING DATE: 2000-05-10
: PRIOR APPLICATION NUMBER: 09/335,409
: PRIOR FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 5
: LENGTH: 7257
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-568-486-5

```

```

Query Match      11.9%: Score 7; DB 4; Length 7257;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      5 SLARAK 11
      |||||||
Db      7250 SLARAK 7256

```

```

RESULT 9
US-09-568-472-5
: Sequence 5, Application US/09568472
: Patent No. 6358719
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/568,472
: CURRENT FILING DATE: 2000-05-10
: PRIOR APPLICATION NUMBER: 09/335,409
: PRIOR FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 5
: LENGTH: 7257
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-568-472-5

```

```

Query Match      11.9%: Score 7; DB 4; Length 7257;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 7250 SLARACK 7256

RESULT 10

US-09-567-899-5

Sequence 5, Application US/09567899

Patent No. 6383787

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James

APPLICANT: Molnar, Istvan

APPLICANT: Zirkle, Ross

APPLICANT: Cyr, Devan

APPLICANT: Goerlach, Joern

TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

FILE REFERENCE: 4-30582A

CURRENT APPLICATION NUMBER: US/09/567,899

CURRENT FILING DATE: 2000-05-10

PRIOR APPLICATION NUMBER: 09/335,409

PRIOR FILING DATE: 1999-06-17

NUMBER OF SEQ ID NOS: 30

SOFTWARE: Patentln Ver. 2.0

SEQ ID NO 5

LENGTH: 7257

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-567-899-5

Query Match 11.9%; Score 7; DB 4; Length 7257;

Best Local Similarity 100.0%; Pred. No. 95;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 SLARACK 11

Db 7250 SLARACK 7256

RESULT 11

US-09-082-358B-24

Sequence 24, Application US/09082358B

Patent No. 6469153

GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.

APPLICANT: Li, Xingquiang

TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,

FILE REFERENCE: 0575/54804

CURRENT APPLICATION NUMBER: US/09/082,358B

CURRENT FILING DATE: 1998-05-20

NUMBER OF SEQ ID NOS: 106

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 24

LENGTH: 15

TYPE: PRT

ORGANISM: murine

US-09-082-358B-24

Query Match 10.2%; Score 6; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKKK 28

Db 5 EKKKKK 10

RESULT 12

US-09-082-358B-62

Sequence 62, Application US/09082358B

Patent No. 6469153

GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.

APPLICANT: Li, Xingquiang

TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,

FILE REFERENCE: 0575/54804

CURRENT APPLICATION NUMBER: US/09/082,358B

CURRENT FILING DATE: 1998-05-20

NUMBER OF SEQ ID NOS: 106

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 62

LENGTH: 15

TYPE: PRT

ORGANISM: murine

US-09-082-358B-62

Query Match 10.2%; Score 6; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKKK 28

Db 10 EKKKKK 15

RESULT 13

US-08-851-843A-201

Sequence 201, Application US/08851843A

Patent No. 6093809

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6093809el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,843A

FILING DATE: 06-MAY-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 201:

SEQUENCE CHARACTERISTICS:

LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-851-843A-201

Query Match 10.2% Score 6; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKKK 28
|||||
DB 8 EKKKKK 13

RESULT 14
US-08-974-549A-320
Sequence 320, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 320:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-320

Query Match 10.2% Score 6; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKKK 28
|||||
DB 8 EKKKKK 13

RESULT 15
US-08-854-050-201
Sequence 201, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 201:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
us-08-854-050-201

Query Match 10.28; Score 6; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 23 EKKKK 28
|||||
Db 8 EKKKK 13

Search completed: January 9, 2003, 12:31:12
Job time: 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:30:29 : Search time 10 seconds
(without alignments)
114.466 Million cell updates/sec

Title: US-09-424-815e-1

Perfect score: 59

Sequence: 1 KVGSLARAGKVRGQTPKVA.....RRFVNVPTFGKKGPNNANS 59

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 118974 seqs, 19401057 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	170	10	US-09-925-301-994
2	28	47.5	118	10	US-09-864-761-35333
3	7	11.9	62	10	US-09-864-761-39484
4	7	11.9	121	10	US-09-764-877-1759
5	5	11.9	227	9	US-09-774-633-155
6	7	11.9	7257	9	US-10-014-717-5
7	6	10.2	14	10	US-09-983-067-1
8	6	10.2	15	9	US-09-945-248-87
9	6	10.2	16	9	US-09-843-676-201
10	6	10.2	17	10	US-09-764-869-729
11	6	10.2	24	10	US-09-864-761-41097
12	6	10.2	28	10	US-09-925-299-1319
13	6	10.2	29	10	US-09-071-838-189
14	6	10.2	31	10	US-09-764-846-248
15	6	10.2	32	10	US-09-925-299-1474
16	6	10.2	36	10	US-09-764-877-1043
17	6	10.2	36	10	US-09-925-300-1820
18	6	10.2	37	10	US-09-925-301-1640
19	6	10.2	43	10	US-09-925-301-1633

20	6	10.2	45	10	US-09-764-869-926	Sequence 926, App
21	6	10.2	48	10	US-09-925-299-1098	Sequence 1098, Ap
22	6	10.2	52	10	US-09-764-869-1228	Sequence 1228, Ap
23	6	10.2	55	10	US-09-764-855-113	Sequence 113, App
24	6	10.2	56	10	US-09-764-847-601	Sequence 601, App
25	6	10.2	57	10	US-09-864-761-47593	Sequence 47593, A
26	6	10.2	59	10	US-09-864-761-47812	Sequence 47812, A
27	6	10.2	60	9	US-09-989-919-118	Sequence 118, App
28	6	10.2	60	10	US-09-925-300-1198	Sequence 1198, Ap
29	6	10.2	62	10	US-09-925-301-1475	Sequence 1475, Ap
30	6	10.2	63	9	US-10-117-604-2	Sequence 2, Appl1
31	6	10.2	63	10	US-09-764-846-260	Sequence 260, App
32	6	10.2	64	10	US-09-864-761-45501	Sequence 45501, A
33	6	10.2	66	10	US-09-764-846-1224	Sequence 1224, App
34	6	10.2	66	10	US-09-764-877-1089	Sequence 1089, App
35	6	10.2	67	10	US-09-864-870-284	Sequence 284, App
36	6	10.2	68	9	US-09-201-936-18	Sequence 18, Appl
37	6	10.2	68	9	US-09-201-936-19	Sequence 19, Appl
38	6	10.2	68	10	US-09-764-846-191	Sequence 191, App
39	6	10.2	69	10	US-09-925-299-1340	Sequence 1340, Ap
40	6	10.2	72	10	US-09-764-846-179	Sequence 179, App
41	6	10.2	72	10	US-09-925-300-1105	Sequence 1105, App
42	6	10.2	73	10	US-09-764-846-242	Sequence 242, App
43	6	10.2	75	10	US-09-833-067-9	Sequence 9, Appl1
44	6	10.2	81	10	US-09-864-761-44564	Sequence 44564, A
45	6	10.2	84	10	US-09-764-846-241	Sequence 241, App

ALIGNMENTS

RESULT 1
US-09-925-301-994
; Sequence 994, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 994
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-994

Query Match 100.0%; Score 59; DB 10; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.4e-47;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKKKKKTKGRARRRQYNNRFVNVPTFGKKGPNNANS 59
|||||
DB 112 KVGSLARAGKVRGQTPKVAKKKKKTKGRARRRQYNNRFVNVPTFGKKGPNNANS 170

RESULT 2
; Sequence 35333, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

```
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 35333
LENGTH: 118
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009704.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 18
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 30
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: SWISSPROT HIT: P35544, EVALUATE 4.00e-23
OTHER INFORMATION: EST_HUMAN HIT: BF347807.1, EVALUATE 6.00e-47
US-09-864-761-35333
```

```
Query Match 47.5% Score 28; DB 10; Length 118;
Best Local Similarity 100.0%; Pred. No. 3.2e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 KVGSLARAGKVRGQTPKVAKEKKKK 28
DB 56 KVGSLARAGKVRGQTPKVAKEKKKK 83
```

```
RESULT 3
US-09-864-761-39484
Sequence 39484, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecomica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39484
LENGTH: 62
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005598.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.7
US-09-864-761-39484
```

```
Query Match 11.9% Score 7; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 23 EKKKKKT 29
|||||
DB 21 EKKKKKT 27

RESULT 4
US-09-764-877-1759
; Sequence 1759, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1759
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-877-1759

Query Match 11.9%; Score 7; DB 10; Length 121;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKKT 29
|||||
DB 114 EKKKKKT 120

RESULT 5
US-09-774-639-155
; Sequence 155, Application US/09774639
; Publication No. US20030003555A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P1
; CURRENT APPLICATION NUMBER: US/09/774,639
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-774-639-155

Query Match 11.9%; Score 7; DB 9; Length 227;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKKT 29
|||||
DB 220 EKKKKKT 226

RESULT 6
US-10-014-717-5
; Sequence 5, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ilgon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Ziklie, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern

; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-10-014-717-5

Query Match 11.9%; Score 7; DB 9; Length 7257;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 SLARAGK 11
|||||
DB 7250 SLARAGK 7256

RESULT 7
US-09-983-067-1
; Sequence 1, Application US/09983067
; Patent No. US20020123101A1
; GENERAL INFORMATION:
; APPLICANT: INOUE, AKIO
; APPLICANT: SHIMIZU, Yoshihiro
; APPLICANT: UEDA, Takuya
; TITLE OF INVENTION: Process for Producing Peptides By Using In Vitro
; FILE REFERENCE: 1752-0151P
; CURRENT APPLICATION NUMBER: US/09/983,067
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: JP 294795/2001
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: JP 227094/2001
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: JP 6910/2001
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: JP 401417/2000
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Bovine sp.
US-09-983-067-1

Query Match 10.2%; Score 6; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKK 28
|||||
DB 9 EKKKK 14

RESULT 8
US-09-945-249-87
; Sequence 87, Application US/09945249
; Patent No. US20020168748A1
; GENERAL INFORMATION:
; APPLICANT: BERLIN, VIVIAN
; APPLICANT: DAMAGNEZ, VERONIQUE
; APPLICANT: SMITH, SUSAN E.
; TITLE OF INVENTION: ASSAYS AND REAGENTS FOR IDENTIFYING ANTI-FUNGAL AGENTS,
; FILE REFERENCE: MIV-074.06
; CURRENT APPLICATION NUMBER: US/09/945,249

; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 09/041,990
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: 08/771,212
; PRIOR FILING DATE: 1996-12-20
; PRIOR APPLICATION NUMBER: 08/631,319
; PRIOR FILING DATE: 1996-04-11
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 87
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide that
; OTHER INFORMATION: corresponds to the C-termini of Phase or Ggrase
; OTHER INFORMATION: substrates
US-09-945-249-87

Query Match 10.2%; Score 6; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28
Db 6 EKKKKK 11

RESULT 9
US-09-843-676-201

; Sequence 201, Application US/09843676
; Patent No. US20020164786A1

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; Lingner, Joachim

; Nakamura, Toru

; Chapman, Karen B.

; Morin, Gregg B.

; Harley, Calvin

; Andrews, William H.

; TITLE OF INVENTION: No. US20020164786A1 Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/843,676

; FILING DATE: 26-Apr-2001

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/854,050

; FILING DATE: 09-MAY-1997

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.

; REGISTRATION NUMBER: 36,429

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 201:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 201:
US-09-843-676-201

Query Match 10.2%; Score 6; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28
Db 8 EKKKKK 13

RESULT 10
US-09-764-869-729

; Sequence 729, Application US/09764869

; Patent No. US20020061521A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007

; CURRENT APPLICATION NUMBER: US/09/764,869

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 729

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-869-729

Query Match 10.2%; Score 6; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28
Db 10 EKKKKK 15

RESULT 11
US-09-864-761-41097

; Sequence 41097, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263,6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41097
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC013746.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.8
; US-09-864-761-41097

Query Match          10.2%  Score 6;  DB 10;  Length 24;
Best Local Similarity 100.0%;  Pred. No. 15;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy 23 EKKKKK 28
    |||||
Db 6 EKKKKK 11

RESULT 12
US-09-925-299-1319
; Sequence 1319, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1319
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (23)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-1319

Query Match          10.2%  Score 6;  DB 10;  Length 28;
Best Local Similarity 100.0%;  Pred. No. 17;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy 24 KKKKKT 29
    |||||
Db 10 KKKKKT 15

RESULT 13
US-09-071-838-189
; Sequence 189, Application US/09071838
; Patent No. US20020152501A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramin
; APPLICANT: Margosian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; TITLE OF INVENTION: Nucleic Acids That Control Seed and
; TITLE OF INVENTION: Fruit Development In Plants
; NUMBER OF SEQUENCES: 324
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,838
; FILING DATE: 01-MAY-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-086100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 189:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-071-838-189

Query Match          10.2%  Score 6;  DB 10;  Length 29;
Best Local Similarity 100.0%;  Pred. No. 17;
Matches 6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy 48 TFGKKK 53
    |||||
Db 1 TFGKKK 6
```

Search completed: January 9, 2003, 12:34:21
Job time : 10 secs

RESULT 14
US-09-764-846-248
; Sequence 248, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT212
; CURRENT APPLICATION NUMBER: US/09/764,846
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 248
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-846-248

Query Match 10.2%; Score 6; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKK 28
|11111
Db 5 EKKKK 10

RESULT 15
US-09-925-299-1474
; Sequence 1474, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1474
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1474

Query Match 10.2%; Score 6; DB 10; Length 32;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 EKKKK 28
|11111
Db 17 EKKKK 22

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:29:54 ; Search time 15 Seconds
(without alignments)
279.052 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59

Sequence: 1 KVHGLARAGKVRGQTPKVA.....RRFNVVPTFGKKGNANS 59

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 263044 seqs, 70945482 residues

Word size : 0

Total number of hits satisfying chosen parameters: 263044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending_Patents_AA_New:*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	54.2	130	5	US-09-513-999C-5650
2	28	47.5	118	6	US-10-203-138A-12417
3	7	11.9	81	6	US-10-302-279-68
4	7	11.9	91	6	US-10-209-582-539
5	6	10.2	16	5	US-09-721-456-320
6	6	10.2	16	6	US-10-325-810-320
7	6	10.2	40	5	US-09-661-665A-2
8	6	10.2	55	5	US-09-513-999C-6360
9	6	10.2	63	5	US-09-513-999C-5320
10	6	10.2	66	1	PCT-US02-32727-3167
11	6	10.2	66	6	US-10-057-498-3167
12	6	10.2	75	1	PCT-US02-32727-18862
13	6	10.2	75	6	US-10-057-498-18862
14	6	10.2	78	5	US-09-724-676-60906
15	6	10.2	78	5	US-09-724-676A-60906
16	6	10.2	79	5	US-09-724-676-90284
17	6	10.2	79	5	US-09-724-676A-90284
18	6	10.2	80	6	US-10-209-582-982
19	6	10.2	81	1	PCT-US02-32727-5391
20	6	10.2	81	6	US-10-057-498-5391
21	6	10.2	88	6	US-10-203-138A-12896
22	6	10.2	89	6	US-10-203-138A-12328
23	6	10.2	90	5	US-09-724-676-84632
24	6	10.2	90	5	US-09-724-676A-84632
25	6	10.2	90	5	US-09-724-676A-84632
26	6	10.2	90	5	US-09-724-676A-84633

27	6	10.2	91	6	US-10-203-138A-11211	Sequence 11211, A
28	6	10.2	94	1	PCT-US02-32727-20914	Sequence 20914, A
29	6	10.2	94	6	US-10-057-498-20914	Sequence 20914, A
30	6	10.2	97	1	PCT-US02-32727-14364	Sequence 14364, A
31	6	10.2	97	1	PCT-US02-32727-29917	Sequence 29917, A
32	6	10.2	97	6	US-10-057-498-14364	Sequence 14364, A
33	6	10.2	99	6	US-10-203-138A-15139	Sequence 15139, A
34	6	10.2	101	5	US-09-513-999C-7902	Sequence 7902, A
35	6	10.2	107	6	US-10-264-237-1659	Sequence 1659, Ap
36	6	10.2	110	6	US-10-276-774-2525	Sequence 2525, Ap
37	6	10.2	114	6	US-10-264-237-2077	Sequence 2077, Ap
38	6	10.2	116	6	US-10-276-774-2213	Sequence 2213, Ap
39	6	10.2	117	6	US-10-264-237-1989	Sequence 1989, Ap
40	6	10.2	125	5	US-09-513-999C-7941	Sequence 7941, Ap
41	6	10.2	131	5	US-09-513-999C-7942	Sequence 7942, Ap
42	6	10.2	132	5	US-09-724-676-84630	Sequence 84630, A
43	6	10.2	132	5	US-09-724-676A-84631	Sequence 84631, A
44	6	10.2	132	5	US-09-724-676A-84630	Sequence 84630, A
45	6	10.2	132	5	US-09-724-676A-84631	Sequence 84631, A

ALIGNMENTS

RESULT 1

US-09-513-999C-5650
; Sequence 5650, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REC
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5650
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 107
; OTHER INFORMATION: Xaa=Lys or Asn
US-09-513-999C-5650

Query Match 54.2%; Score 32; DB 5; Length 130;
Best Local Similarity 100.0%; Pred. No. 6.5e-26;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVHGLARAGKVRGQTPKVAQEKKKKTGRA 32
|||||
DB 75 KVHGLARAGKVRGQTPKVAQEKKKKTGRA 106
|||||

RESULT 2

US-10-203-138A-12417
; Sequence 12417, Application US/10203138A
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 8
; CURRENT APPLICATION NUMBER: US/10/203,138A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 15438
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 12417
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009704.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 81
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P35544, EVALUE 4.00e-23
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF347807.1, EVALUE 6.00e-47
US-10-203-138A-12417

Query Match 47.5%; Score 28; DB 6; Length 118;
Best Local Similarity 100.0%; Pred. No. 8.5e-22;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLRACKVRGQTPKVAQEKKKK 28
DB 56 KVGSLRACKVRGQTPKVAQEKKKK 83

RESULT 3
US-10-302-279-68
; Sequence 68, Application US/10302279
; GENERAL INFORMATION:
; APPLICANT: Dean, Michael Carlton
; Hahn, Heidi Eve
; Wicking, Carol
; Christiansen, Jeffrey G.
; Zaphiropoulos, Peter G.
; Gallani, Mae R.
; Shanley, Susan Mary
; Chidambaram, Abirami
; Vorechovsky, Igor
; Holmberg-Lindstrom, Erika
; TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/302,279
; FILING DATE: 22-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/857,636
; FILING DATE: 16-MAY-1997

; APPLICATION NUMBER: US 60/017,906
; FILING DATE: 17-MAY-1996
; APPLICATION NUMBER: AU P00011
; FILING DATE: 21-MAY-1996
; APPLICATION NUMBER: AU P00363
; FILING DATE: 07-JUN-1996
; APPLICATION NUMBER: US 60/019,765
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Laurence J.
; REGISTRATION NUMBER: 35, 551
; REFERENCE/DOCKET NUMBER: 015280-27820005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 81 amino acid
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-10-302-279-68

Query Match 11.9%; Score 7; DB 6; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKKT 29
DB 31 EKKKKKT 37

RESULT 4
US-10-209-582-539
; Sequence 539, Application US/10209582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM009CIN
; CURRENT APPLICATION NUMBER: US/10/209,582
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 09/758,461
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1030
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 539
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-582-539

Query Match 11.9%; Score 7; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KKKKKTG 30
DB 37 KKKKKTG 43

RESULT 5
US-09-721-456-320
; Sequence 320, Application US/09721456
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru

; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/721,456
; FILING DATE: 22-Nov-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-Nov-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 320:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 320:
US-09-721-456-320

Query Match 10.2%; Score 6; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28

Db 8 EKKKKK 13

RESULT 6

US-10-325-810-320

; Sequence 320, Application US/10325810
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin B.
; Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 633
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/325,810
; FILING DATE: 20-Dec-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/402,181
; FILING DATE: 29-Sep-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ausenhus, Scott L.
; REGISTRATION NUMBER: 42,271
; REFERENCE/DOCKET NUMBER: 015389-002620US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 320:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 320:
US-10-325-810-320

Query Match 10.2%; Score 6; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28

Db 8 EKKKKK 13

```
RESULT 7
US-09-661-669A-2
; Sequence 2, Application US/09661669A
; GENERAL INFORMATION:
; APPLICANT: Ferullo, Jean-Marc
; APPLICANT: Paget, Eric
; TITLE OF INVENTION: Hydro-Phenyl Pyruvate Dioxigenase Fused to a Peptide
; TITLE OF INVENTION: Signal, the DNA Sequence and Obtaining Plants
; TITLE OF INVENTION: containing such gene, Tolerant to Herbicides
; FILE REFERENCE: 5500*51
; CURRENT APPLICATION NUMBER: US/09/661,669A
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 40
; TYPE: PRT
; ORGANISM: synthetic construct
US-09-661-669A-2

Query Match          10.2%; Score 6; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKK 28
Db 14 EKKKK 19

RESULT 8
US-09-513-999C-6360
; Sequence 6360, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6360
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 8
; OTHER INFORMATION: Xaa-Leu or Pro
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 11
; OTHER INFORMATION: Xaa-Glu or Lys
US-09-513-999C-6360

Query Match          10.2%; Score 6; DB 5; Length 55;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKK 28
Db 31 EKKKK 36

RESULT 9
US-09-513-999C-5320
; Sequence 5320, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
```

```
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5320
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5320

Query Match          10.2%; Score 6; DB 5; Length 63;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKK 28
Db 47 EKKKK 52

RESULT 10
PCT-US02-32727-3167
; Sequence 3167, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siging
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 3167
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-3167

Query Match          10.2%; Score 6; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GSLARA 9
Db 13 GSLARA 18

RESULT 11
US-10-057-498-3167
; Sequence 3167, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
```

; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 3167
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-3167

Query Match 10.2%; Score 6; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GSLARA 9
Db 13 GSLARA 18
|||||

RESULT 12
PCT-US02-32727-18862
; Sequence 18862, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylan
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Derrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30952
; SEQ ID NO 18862
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-18862

Query Match 10.2%; Score 6; DB 1; Length 75;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VHGSALA 7
Db 38 VHGSALA 43
|||||

RESULT 13
US-10-057-498-18862
; Sequence 18862, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 18862
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-18862

Query Match 10.2%; Score 6; DB 6; Length 75;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VHGSALA 7
Db 38 VHGSALA 43
|||||

RESULT 14
US-09-724-676-60906
; Sequence 60906, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60906
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-60906

Query Match 10.2%; Score 6; DB 5; Length 78;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28
Db 59 EKKKKK 64
|||||

RESULT 15
US-09-724-676A-60906
; Sequence 60906, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60906
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-60906

Query Match 10.2%; Score 6; DB 5; Length 78;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKK 28
Db 59 EKKKKK 64
|||||

Search completed: January 9, 2003, 12:34:04
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:29:39 ; Search time 143 Seconds
(without alignments)
266.009 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59

Sequence: 1 KVHGLARAGKRGQTPKVA.....RRFVNVVPTFGKKKGNANS 59

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 4569144 seqs, 644733110 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending_Patents_AA_Main.*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	59	18	US-09-424-815-1
2	59	100.0	59	18	US-09-424-815C-1
3	59	100.0	59	18	US-09-424-815E-1
4	59	100.0	59	21	US-09-732-210-1596
5	59	100.0	59	27	US-60-169-340-1596
6	59	100.0	63	22	US-09-834-366-24153

7	59	100.0	63	22	US-09-834-366-24373	Sequence 24373, A
8	59	100.0	63	27	US-60-197-873-24153	Sequence 24153, A
9	59	100.0	63	27	US-60-197-873-24373	Sequence 24373, A
10	59	100.0	66	1	PCT-US01-18569-3540	Sequence 3540, Ap
11	59	100.0	66	26	US-10-284-049-3540	Sequence 3540, Ap
12	59	100.0	133	21	US-09-791-537-25338	Sequence 25338, A
13	59	100.0	133	21	US-09-791-537-31129	Sequence 31129, A
14	59	100.0	133	21	US-09-791-537-38099	Sequence 38099, A
15	59	100.0	133	21	US-09-791-537-100293	Sequence 100293, A
16	59	100.0	133	21	US-09-791-537-102300	Sequence 102300, A
17	59	100.0	142	21	US-09-760-469-1748	Sequence 1748, Ap
18	59	100.0	142	22	US-09-874-043-619	Sequence 619, App
19	59	100.0	142	22	US-09-875-195-619	Sequence 619, App
20	59	100.0	142	26	US-10-216-583-1748	Sequence 1748, Ap
21	59	100.0	145	21	US-09-760-469-1288	Sequence 1288, Ap
22	59	100.0	145	26	US-10-216-583-1288	Sequence 1288, Ap
23	59	100.0	170	1	PCT-US00-05882-994	Sequence 994, App
24	59	100.0	170	21	US-09-760-469-1744	Sequence 1744, Ap
25	59	100.0	170	23	US-09-925-301-994	Sequence 994, App
26	59	100.0	170	26	US-10-216-583-1744	Sequence 1744, Ap
27	59	100.0	173	21	US-09-760-469-1284	Sequence 1284, Ap
28	59	100.0	173	26	US-10-216-583-1284	Sequence 1284, Ap
29	54	91.5	267	27	US-60-340-187-651	Sequence 651, App
30	54	91.5	337	27	US-60-340-187-668	Sequence 668, App
31	41	69.5	46	1	PCT-US01-14827-10440	Sequence 10440, A
32	39	66.1	108	1	PCT-US01-08631-57587	Sequence 57587, A
33	39	66.1	108	1	PCT-US01-08631-57878	Sequence 57878, A
34	37	62.7	157	1	PCT-US00-26524B-6729	Sequence 6729, Ap
35	37	62.7	157	25	US-10-106-698-6739	Sequence 6739, Ap
36	32	54.2	312	1	PCT-US01-08656-7357	Sequence 7357, Ap
37	32	54.2	312	1	PCT-US01-14827-10441	Sequence 10441, A
38	31	52.5	61	1	PCT-US01-08631-57589	Sequence 57589, A
39	31	52.5	90	1	PCT-US01-08631-57879	Sequence 57879, A
40	28	47.5	118	1	PCT-US01-00663-28387	Sequence 28387, A
41	28	47.5	118	22	US-09-864-761-35333	Sequence 35333, A
42	28	47.5	118	25	US-10-182-993-27520	Sequence 27520, A
43	28	47.5	118	25	US-10-182-995-21805	Sequence 21805, A
44	28	47.5	118	25	US-10-182-997-20446	Sequence 20446, A
45	28	47.5	118	25	US-10-182-998-12104	Sequence 12104, A

ALIGNMENTS

RESULT 1
US-09-424-815-1
; Sequence 1, Application US/09424815
; GENERAL INFORMATION:
; APPLICANT: Nibbeling, Petrus Hendricus
; APPLICANT: Hiemstra, Pieter Sicco
; APPLICANT: Van den Barselaar, Maria Theodora
; APPLICANT: Pauwels, Ernest Karl Jacob
; APPLICANT: Feitsma, Rolf Ide Johannes
; TITLE OF INVENTION: Antimicrobial Peptides Derived From Ubiquitin
; FILE REFERENCE: 702 991768
; CURRENT APPLICATION NUMBER: US/09/424,815
; CURRENT FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/NL98/00311
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: NL 1006164
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Microsoft Word 97 SR-2
; SEQ ID NO 1
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Mammalian
US-09-424-815-1

Query Match 100.0%; Score 59; DB 18; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.2e-46;

; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 24153
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-24153

Query Match 100.0%; Score 59; DB 22; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 59
|||||
Db 5 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 63

RESULT 7

US-09-834-366-24373
; Sequence 24373, Application US/09834366
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 24373
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-24373

Query Match 100.0%; Score 59; DB 22; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 59
|||||
Db 5 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 63

RESULT 8

US-09-197-873-24153
; Sequence 24153, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm
; SEQ ID NO 24153
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-197-873-24153

Query Match 100.0%; Score 59; DB 27; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 59
|||||
Db 5 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 63

RESULT 9

US-60-197-873-24373
; Sequence 24373, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; APPLICANT: Giordano, Jean-Yves
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 24373
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-197-873-24373

Query Match 100.0%; Score 59; DB 27; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 59
|||||
Db 5 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 63

RESULT 10

PCT-US01-18569-3540
; Sequence 3540, Application PC/TUS0118569
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAI33PCT
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3540
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-18569-3540

Query Match 100.0%; Score 59; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 59
|||||
Db 8 KVGSLARAGKVRGQTPKVAKEKKKKTKGRAKRRMOYNNRRFVNVVPTFGKKKGNPNS 66

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31129
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-31129

```

```
Query Match      100.0%; Score 59; DB 21; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 14
US-09-791-537-38099
; Sequence 38099, Application US/09791537

```
Query Match      100.0%; Score 59; DB 21; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 15
US-09-791-537-100293
; Sequence 100293, Application US/09791537

```
Query Match      100.0%; Score 59; DB 21; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 11
US-10-264-049-3540
; Sequence 3540, Application US/10264049
: GENERAL INFORMATION:

```
Query Match      100.0%; Score 59; DB 26; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 12
US-09-791-537-25338
: Sequence 25338. Application US/09791537

```
Query Match      100.0%; Score 59; DB 21; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-46;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 13
US-09-791-537-31129
: Sequence 31129, Application US/09791537

Search completed: January 9, 2003, 12:33:43
Job time : 144 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:28:38 : Search time 19 seconds
(without alignments)
296.523 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59

Sequence: 1 KVGSLARAGKVRGQTPKVA.....RRFVNVPTFGKKGPNNNS 59

Scoring table:

GAPOP 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: p1r3:*
2: p1r1:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	133	1 A47416	ubiquitin-like pro
2	59	100.0	133	2 JCI278	ubiquitin-like pro
3	59	100.0	133	3 146346	ribosomal protein
4	22	37.3	62	2 H85342	RIBOSOMAL PROTEIN
5	22	37.3	68	2 F84580	40S ribosomal prot
6	20	33.9	229	2 S38383	SEB4B protein - hu
7	13	23.0	130	2 T15642	hypothetical prote
8	12	20.3	58	2 A71604	ribosomal protein
9	12	20.3	61	2 T39834	ribosomal protein
10	12	20.3	63	1 S67074	ribosomal protein
11	8	13.6	135	2 A97094	uncharacterized pr
12	7	11.9	134	2 S02776	DNA-binding protei
13	7	11.9	154	2 B75305	hypothetical prote
14	7	11.9	199	2 E84782	hypothetical prote
15	7	11.9	215	2 I52523	nucleoporin p62 ho
16	7	11.9	233	2 H86318	probable 60S ribos
17	7	11.9	233	2 D96768	protein 60S riboso
18	7	11.9	233	2 E96768	protein 60S riboso
19	7	11.9	283	2 T47174	hypothetical prote
20	7	11.9	314	2 D70313	riboflavin kinase
21	7	11.9	346	2 T19629	hypothetical prote
22	7	11.9	424	2 T36154	probable ABC-type
23	7	11.9	454	2 E84153	replicative DNA he
24	7	11.9	710	2 T49516	Atu related protei
25	7	11.9	863	2 F85343	hypothetical prote
26	7	11.9	1234	1 J01979	structural polypor
27	7	11.9	2539	2 B71619	hypothetical prote
28	6	10.2	46	2 T28799	hypothetical prote
29	6	10.2	67	2 B69830	hypothetical prote

30	6	10.2	70	2 H97135	uncharacterized ph
31	6	10.2	74	2 T47434	hypothetical prote
32	6	10.2	80	2 B82267	exodeoxyribonucle
33	6	10.2	81	2 S66013	ribosomal protein
34	6	10.2	92	2 T36134	hypothetical prote
35	6	10.2	92	2 H82867	transcription regu
36	6	10.2	98	2 H96585	hypothetical prote
37	6	10.2	102	2 T07078	cold stress protei
38	6	10.2	104	2 C87604	vitulence associat
39	6	10.2	105	2 PW0017	hypothetical prote
40	6	10.2	105	2 PW0018	hypothetical prote
41	6	10.2	106	2 S59536	cold stress protei
42	6	10.2	115	2 S69849	hypothetical prote
43	6	10.2	117	2 H84651	hypothetical prote
44	6	10.2	123	2 B71268	probable ribosomal
45	6	10.2	125	1 R3RT25	ribosomal protein

ALIGNMENTS

```

RESULT 1
A47416
ubiquitin-like protein / ribosomal protein S30, cytosolic [validated] - rat
N:Contains: ribosomal protein S30; ubiquitin-like protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jul-2000 #sequence:revision 21-Jul-2000 #text:change 21-Jul-2000
C:Accession: A47416; B47416; S18101
J:Olvera, J.; Wool, I.G.
J: Biol. Chem. 268, 17967-17974, 1993
A:Title: The carboxyl extension of a ubiquitin-like protein is rat ribosomal protein
A:Reference number: A47416; MUID:93352612; PMID:8394356
A:Accession: A47416
A:Molecule type: mRNA
A:Residues: 1-133 <OLV>
A:Cross-references: EMBL:X62671; NID:g407165; PIDN:CAA44545.1; PID:g57566
A:Accession: B47416
A:Molecule type: protein
A:Residues: 75-92 <OL2>
A>Note: the proteins are designated as ubiquitin-like protein and ribosomal protein S
C:Superfamily: ubiquitin-like protein / rat ribosomal protein S30; ubiquitin homology
C:Keywords: protein biosynthesis; ribosome
F:1-74/Product: ubiquitin-like protein [status predicted <UBI>
F:1-74/Domain: ubiquitin homology <UBH>
F:75-133/Product: ribosomal protein S30 [status experimental <RIB>

Query Match          100.0%; Score 59; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 5.2e-52;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVGSLARAGKVRGQTPKVAKKKKKTRARRRQYNNRFVNVPTFGKKGPNNNS 59
DB 75 KVGSLARAGKVRGQTPKVAKKKKKTRARRRQYNNRFVNVPTFGKKGPNNNS 133

RESULT 2
JCI278
ubiquitin-like protein / ribosomal protein S30, cytosolic - human
N:Alternate names: fau protein
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence:revision 30-Sep-1993 #text:change 24-Sep-1999
C:Accession: JCI278; I37387; S68911; S21445; S21450
R:Kas, K.; Michiels, L.; Merregaert, J.
Blochem. Biophys. Res. Commun. 187, 927-933, 1992
A:Title: Genomic structure and expression of the human fau gene: Encoding the ribosome
A:Reference number: JCI278; MUID:92412144; PMID:1326960
A:Accession: JCI278
A:Molecule type: DNA
A:Residues: 1-133 <KAS>
A:Cross-references: EMBL:X65921; NID:g31304; PIDN:CAA46714.1; PID:g31305
R:Michiels, L.; Van der Raaij-MacLeod, E.; Van Hasselt, F.; Kas, K.; Merregaert, J.
Oncogene 8, 2537-2546, 1993
A:Title: fau cDNA encodes a ubiquitin-like-S30 fusion protein and is expressed as an

```

A:Reference number: I37387; MUID:93368957; PMID:8395663
 A:Accession: I37387
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-133 <RES>
 A:Cross-references: EMBL:X65923; NID:g31302; PIDN:CAA6716.1; PID:g31303
 R:Vladimirov, S.N.; Ivanov, A.V.; Karpova, G.G.; Musolyanov, A.K.; Egorov, T.A.; Thiede, R.J. Biochem. 239, 144-149, 1996
 A:Title: Characterization of the human small-ribosomal-subunit proteins by N-terminal an
 A:Reference number: S68911; MUID:96305378; PMID:8706699
 A:Accession: S68911
 A:Molecule type: protein
 A:Residues: 75-99 <VLA>
 A:Gene: fau
 C:Genetics:
 A:Introns: 25/3; 74/1; 92/3
 A:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
 C:Keywords: protein biosynthesis; ribosome
 F:1-74/Domain: ubiquitin homology <DBH>
 F:75-133/Product: ribosomal protein S30, cytosolic #status experimental <MAT>

Query Match 100.0%; Score 59; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 5.2e-52;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHGSLARACKVGRGTPKVAKEKKKKTKGRAKRMQYNNRFVNVVPTFGKKGPNNNS 59
 Db 75 KKHGSLARACKVGRGTPKVAKEKKKKTKGRAKRMQYNNRFVNVVPTFGKKGPNNNS 133

RESULT 3

I48346
 ribosomal protein fau - mouse
 N:Alternate names: gene fau protein; monoclonal nonspecific suppressor factor beta
 C:Species: Mus musculus (house mouse)
 C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 24-Sep-1999
 C:Accession: I48346; A56532; I59368; S21452
 R:Michlels, L.; Van der Raaij-MacLeod, E.; Van Hasselt, F.; Kas, K.; Merregaert, J.
 Oncogene 8, 2537-2546, 1993

A:Title: fau cDNA encodes a ubiquitin-like-S30 fusion protein and is expressed as an ant
 A:Reference number: I37387; MUID:93368957; PMID:8395663
 A:Accession: I48346
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-133 <RES>
 A:Cross-references: EMBL:X65922; NID:g50949; PIDN:CAA6715.1; PID:g50950
 R:Castreels, D.; Polter, C.; Guenet, J.L.; Merregaert, J.
 Genomics 25, 291-294, 1995
 A:Title: The mouse fau gene: genomic structure, chromosomal localization, and characteri
 A:Reference number: A56532; MUID:95293388; PMID:7774934
 A:Accession: A56532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-133 <CAS>

A:Cross-references: GB:L37315; NID:g497610; PIDN:AA91564.1; PID:g497611
 A:Note: authors translated the codon GTC for residue 119 as Arg, and GTC for residue 120
 R:Nakamura, M.; Xavier, R.M.; Tsunematsu, T.; Tanigawa, Y.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3463-3467, 1995
 A:Title: Molecular cloning and characterization of a cDNA encoding monoclonal nonspecific
 A:Reference number: I59368; MUID:95241522; PMID:7724584
 A:Accession: I59368
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-133 <RE2>
 A:Cross-references: GB:D26610; NID:g1060926; PIDN:BAA05655.1; PID:g1060927
 C:Genetics:
 A:Gene: fau
 A:Introns: 25/3; 74/1; 92/3
 C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
 F:1-74/Domain: ubiquitin homology <DBH>

Query Match 100.0%; Score 59; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 5.2e-52;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KKHGSLARACKVGRGTPKVAKEKKKKTKGRAKRMQYNNRFVNVVPTFGKKGPNNNS 59
 Db 75 KKHGSLARACKVGRGTPKVAKEKKKKTKGRAKRMQYNNRFVNVVPTFGKKGPNNNS 133

RESULT 4

H85342
 RIBOSOMAL PROTEIN S30 homolog [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: H85342
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold sp
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488; PMID:10617198
 A:Accession: H85342
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-62 <STO>

A:Cross-references: GB:NC_001268; NID:g7269837; PIDN:CAB79697.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: Atg29390
 A:Map position: 4

Query Match 37.3%; Score 22; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 4.3e-15;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHGSLARACKVGRGTPKVAQ 22
 Db 3 KKHGSLARACKVGRGTPKVAQ 24

RESULT 5

F84580
 40S ribosomal protein S30 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
 C:Accession: F84580
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.T.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84580
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-68 <STO>

A:Cross-references: GB:AE002093; NID:g3687243; PIDN:AAC62141.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: Atg2919750
 A:Map position: 2
 C:Superfamily: yeast ribosomal protein S30_e

Query Match 37.3%; Score 22; DB 2; Length 68;
 Best Local Similarity 100.0%; Pred. No. 4.7e-15;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHGSLARACKVGRGTPKVAQ 22
 Db 9 KKHGSLARACKVGRGTPKVAQ 30

RESULT 6

S38383
 SEB4B protein - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 08-Sep-1997
 C:Accession: S38383
 R:Ruehlmann, A.; Gupta, A.; Terworst, C.

submitted to the EMBL Data Library, September 1993
A:Description: A novel murine RRM-type protein and its human homology.
A:Reference number: S38382

A:Accession: S38383
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <RUE>
A:Cross-references: EMBL:X75315
C:Superfamily: unassigned ribonucleoprotein repeat homology <RRM1>
F:25-91/domain: ribonucleoprotein repeat homology <RRM1>

Query Match 33.9%; Score 20; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 35 RMQYRRRVNVPYTGKKK 54
Db 1 RMQYRRRVNVPYTGKKK 20

RESULT 7

T15642
hypothetical protein C26F1.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C:Accession: T15642
R:Geisel, C.; Stellyes, L.; Bradshaw, H.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of *C. elegans* cosmid C26F1.
A:Reference number: Z18381
A:Accession: T15642
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-130 <GER>
A:Cross-references: EMBL:U53148; NID:g1255375; PID:g1255381; PIDN:AAE37076.1; GSPDB:GN0C
C:Experimental source: strain Bristol N2; clone C26F1
C:Genetics:
A:Gene: CESP:C26F1.4
A:Map position: 5
A:Introns: 27/3; 71/1
C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 22.0%; Score 13; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVHGSILARAGKVR 13
Db 72 KVHGSILARAGKVR 84

RESULT 8

A71604
ribosomal protein S30 PFB0885v - malaria parasite (*Plasmodium falciparum*)
C:Species: *Plasmodium falciparum*
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 13-Aug-1999
R:Gardner, M.J.; Petrelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: A71604
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-58 <GAR>
A:Cross-references: GB:AE001422; GB:AE001362; NID:g3845298; PIDN:AACT1966.1; PID:g384530
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0885v
C:Superfamily: yeast ribosomal protein S30.e

Query Match 20.3%; Score 12; DB 2; Length 58;

Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVHGSILARAGK 12
Db 3 KVHGSILARAGK 14

RESULT 9

T39834
ribosomal protein s30 - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Jun-2000
C:Accession: T39834
R:Layne, M.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21884
A:Accession: T39834
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-61 <LYN>
A:Cross-references: EMBL:AL021839; PIDN:CAI17057.2; GSPDB:GN00067; SPDB:SPBC1967.03c
A:Experimental source: strain 972h-; cosmid c1967
C:Genetics:
A:Gene: SPDB:SPBC1967.03c
A:Map position: 2
A:Introns: 13/3; 56/3
C:Superfamily: yeast ribosomal protein S30.e

Query Match 20.3%; Score 12; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KVHGSILARAGK 12
Db 3 KVHGSILARAGK 14

RESULT 10

S67074
ribosomal protein S30.e, cytosolic - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein O4725; protein YLR287c-a; protein YOR182c
C:Species: *Saccharomyces cerevisiae*
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: S67074; S70775; S70776; S70774
R:Hughes, B.; Pohl, T.M.
submitted to the Protein Sequence Database, Julv 1996
A:Reference number: S66885
A:Accession: S67074
A:Molecule type: DNA
A:Residues: 1-63 <HUG>
A:Cross-references: EMBL:Z75090; GSPDB:GN00015; MIPS:YOR182c; NID:g1420438; PIDN:CAA9
A:Experimental source: strain S288C
A:Genetics: YS30B
J:Baker, R.T.; Williamson, N.A.; Wettenthal, R.E.H.
J. Biol. Chem. 271, 13549-13555, 1996
A:Title: The yeast homolog of mammalian ribosomal protein S30 is expressed from a dup
A:Reference number: S70774; MUID:96278780; PMID:8662789
A:Accession: S70775
A:Molecule type: DNA
A:Residues: 1-63 <BAK>
A:Cross-references: EMBL:U48700; NID:g1256752; PIDN:AAAC49317.1; PID:g1256753
A:Genetics: YS30A
A:Accession: S70776
A:Molecule type: mRNA
A:Residues: 1-63 <BAW>
A:Cross-references: EMBL:U48699; NID:g1256750; PIDN:AAAC49316.1; PID:g1256751
A:Genetics: YS30A
A:Accession: S70774
A:Molecule type: protein
A:Residues: 2-63 <BA3>
C:Genetics: <YS30B>
A:Gene: SGD:RPS30B; MIPS:YOR182c

A:Cross-references: MIPS:YOR182c; SGD:S0005708
A:Map position: 15R
A:Introns: 1/3
C:Genetics: <YXS30A>
A:Gene: SGD:RPS30A; MIPS:YLR287c-a
A:Cross-references: MIPS:YLR287c-a; SGD:S0004278
A:Map position: 12R
A:Introns: 1/3
C:Superfamily: yeast ribosomal protein S30_e
C:Keywords: cytosol; protein biosynthesis; ribosome
F:2-63/Product: ribosomal protein S30_e, cytosolic #status experimental <KAT>

Query Match 20.3%; Score 12; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 4.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVHGSILARAGKV 12
DB 3 KVHGSILARAGKV 14

RESULT 11
A97094
uncharacterized protein, YUDF B. subtilis ortholog [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: A97094
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: A97094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-135 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79540.1; PID:g15024526; GSPDB:GNO0168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1573

Query Match 13.6%; Score 8; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 KOEKKKKK 28
DB 124 KOEKKKKK 131

RESULT 12
S02776
DNA-binding protein H-NS - Proteus vulgaris
C:Species: Proteus vulgaris
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 13-Sep-1998
C:Accession: S02776
R:la Teana, A.; Falconi, M.; Scarlato, V.; Lammi, M.; Pon, C.L.
FEBS Lett. 244, 34-38, 1989
A:Title: Characterization of the structural genes for the DNA-binding protein H-NS in E
A:Reference number: S02775; MUID:89171270; PMID:2494066
A:Accession: S02776
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-134 <LAL>
C:Genetics:
A:Gene: hns
C:Function:
A:Description: binds double-stranded DNA; influences regulation of gene expression at th
C:Superfamily: DNA-binding protein H-NS
C:Keywords: DNA binding; transcription regulation

Query Match 11.9%; Score 7; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 9.8;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KTGRAKR 34
DB 83 KTGRAKR 89

RESULT 13
B75305
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: B75305
R:White, O.; Eisen, J.A.; Hefelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75305
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <WHI>
A:Cross-references: GB:AE002051; GB:AE000513; NID:g6459976; PIDN:AAF11732.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2179
A:Map position: 1

Query Match 11.9%; Score 7; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RAKVRG 14
DB 39 RAKVRG 45

RESULT 14
E84782
probable homeodomain transcription factor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84782
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bento, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-199 <STO>
A:Cross-references: GB:AE002093; NID:g4415906; PIDN:AAD20137.1; GSPDB:GNO0139
C:Genetics:
A:Gene: At2g36610
A:Map position: 2

Query Match 11.9%; Score 7; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 OEKKKKK 28
DB 68 OEKKKKK 74

RESULT 15
I52523
nucleoporin p62 homolog - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 05-Nov-1999

C:Accession: I52523
R:Wang, Z.O.; Akmal, K.M.; Kim, K.H.
Biol. Reprod. 51, 1022-1030, 1994
A:Title: An unusual nucleoporin-related messenger ribonucleic acid is present in the ger
A:Reference number: I52523; MUID:95151924; PMID:7849178
A:Accession: I52523
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1215 <RES>
A:Cross-references: GB:S75997; NID:9913245; PIDN:AB33384.1; PID:9913246
A:Experimental source: testis

Query Match 11.9%; Score 7; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 KKKKKTG 30
|||||
DB 52 KKKKKTG 58

Search completed: January 9, 2003, 12:30:51
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:25:48 : Search time 10 seconds
(without alignments)
244.710 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59
Sequence: 1 KVGSLRACKVRCQTPKVA.....RRFVNVVPTGKKKGPANNS 59

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40:*

| Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	59	1	Q05472 homo sapien
2	22	37.3	62	1	RS30_HUMAN
3	18	30.5	59	1	RS30_ARATH
4	12	20.3	58	1	RS30_ORYLA
5	12	20.3	61	1	RS30_PLAFA
6	12	20.3	61	1	RS30_SCHPO
7	11	18.6	62	1	RS30_YEAST
8	7	11.9	133	1	RS30_ONCMY
9	7	11.9	133	1	HNS_PROVU
10	7	11.9	489	1	YOC2_CAEEL
11	7	11.9	1254	1	COLB_HUMAN
12	6	10.2	78	1	POLS_EEYVM
13	6	10.2	78	1	R27A_DICDI
14	6	10.2	80	1	RS18_BACSU
15	6	10.2	105	1	EX75_VIBCH
16	6	10.2	105	1	YGRM_MICEC
17	6	10.2	105	1	YGRM_MICRO
18	6	10.2	123	1	RL19_TREPA
19	6	10.2	123	1	RS25_HUMAN
20	6	10.2	127	1	RS11_RICPR
21	6	10.2	128	1	Y082_CAEEL
22	6	10.2	130	1	YNO3_YEAST
23	6	10.2	137	1	YKDI_YEAST
24	6	10.2	142	1	RS16_YEAST
25	6	10.2	142	1	RS16_YEAST
26	6	10.2	157	1	END7_BP14
27	6	10.2	167	1	YK20_YEAST
28	6	10.2	175	1	RY15_LEIDO
29	6	10.2	178	1	FYV5_MOUSE
30	6	10.2	188	1	CC42_CAEEL
31	6	10.2	188	1	RASK_MEICA
32	6	10.2	195	1	PAPH_ECOLI
33	6	10.2	195	1	PRSH_ECOLI
34	6	10.2	198	1	RMRP_YEAST

34	6	10.2	208	1	RASM_HUMAN	O14807 homo sapien
35	6	10.2	208	1	RASM_MOUSE	O08989 mus musculus
36	6	10.2	208	1	RASM_RAT	P97538 rattus norv
37	6	10.2	208	1	RISA_BUCAI	P57212 buchnera ap
38	6	10.2	209	1	RHO1_YEAST	P06780 saccharomyc
39	6	10.2	211	1	T402_BURCE	P24536 burkholderi
40	6	10.2	217	1	KS1_HYDAT	P38978 hydra atten
41	6	10.2	219	1	PLIZ_BACSU	P35536 bacillus su
42	6	10.2	220	1	UNG_BUCAI	P57280 buchnera ap
43	6	10.2	221	1	RS3A_PYRAE	O82121 pyrobaculum
44	6	10.2	233	1	Y008_YEAST	P47006 saccharomyc
45	6	10.2	235	1	CAV1_CAEEL	O94051 caenorhabdi

ALIGNMENTS

```

RESULT 1
ID      RS30_HUMAN          STANDARD:      PRT:      59 AA.
AC      005472: Q05261;
DT      01-JUN-1994 (rel. 29, Created)
DT      01-JUN-1994 (rel. 29, Last sequence update)
DT      16-OCT-2001 (rel. 40, Last annotation update)
DE      40S ribosomal protein S30.
GN      PAU.
OS      Homo sapiens (Human),
OS      Mus musculus (Mouse),
OS      Rattus norvegicus (Rat),
OS      Cricetulus griseus (Chinese hamster), and
OS      Sus scrofa (pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
OX      NCBI_TaxID=9606, 10090, 10116, 10029, 9823;
NM      [1]
RP      SEQUENCE FROM N.A.
RC      SPECIES=Human:
RX      MEDLINE=92412144; PubMed=1326960;
RA      Kas K., Michiels L., Merregaert J.;
RT      "Genomic structure and expression of the human fau gene: encoding the
RL      ribosomal protein S30 fused to a ubiquitin-like protein.";
RN      Blochem. Biophys. Res. Commun. 187:927-933(1992).
[2]
RP      SEQUENCE FROM N.A.
RC      SPECIES=Human, and Mouse;
RX      MEDLINE=93368957; PubMed=8395683;
RA      Michiels L., van der Raaij-van der Wal A.C., van Hasselt F., Kas K.,
RT      Merregaert J.;
RT      "fau cDNA encodes a ubiquitin-like-S30 fusion protein and is
RT      expressed as an antisense sequence in the Finkel-Biskis-Reilly murine
RL      sarcoma virus.";
RN      Oncogene 8:2537-2546(1993).
[3]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.
RC      SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Liver;
RX      MEDLINE=93352612; PubMed=8394356;
RA      Olvera J., Wool I.G.;
RT      "The carboxyl extension of a ubiquitin-like protein is rat ribosomal
RT      protein S30.";
RN      J. Biol. Chem. 268:17967-17974(1993).
[4]
RP      SEQUENCE FROM N.A.
RC      SPECIES=Mouse; STRAIN=BALB/c;
RX      MEDLINE=95241522; PubMed=7724584;
RA      Nakamatsu M., Xavier R.M., Tsunematsu T., Tanigawa Y.;
RT      "Molecular cloning and characterization of a cDNA encoding monoclonal
RT      non-specific suppressor factor.";
RN      Proc. Natl. Acad. Sci. U.S.A. 92:3463-3467(1995).
[5]
RP      SEQUENCE FROM N.A.
RC      SPECIES=Mouse; STRAIN=BALB/c; TISSUE=Liver;
RX      MEDLINE=95293388; PubMed=7774934;
RA      Casteels D., Pollier C., Guenet J.-L., Merregaert J.;

```

RT "The mouse Fau gene: genomic structure, chromosomal localization, and
RT characterization of two retrospodogenes.";
RT Genomics 25:291-294(1995).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=C.griusens;
RA Rossman T.G., Wang Z.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=PIG; TISSUE=uterus;
RA MEDLINE-9724440; PubMed-9089280;
RX Chwetzoff S., D'Andrea S.;
RT "ubiquitin is physiologically induced by interferons in luminal
RT epithelium of porcine uterine endometrium in early pregnancy: global
RT RT-PCR cDNA in place of RNA for differential display screening";
RL FEBS Lett. 405:148-152(1997).
CC -1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
CC C-TERMINAL EXTENSION PROTEIN (CEP) OF A UBIQUITIN-LIKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65921; CAA46714.1; ALT_INIT.
DR EMBL; X65923; CAA46716.1; ALT_INIT.
DR EMBL; X62671; CAA44546.1; -.
DR EMBL; X62671; CAA44545.1; ALT_INIT.
DR EMBL; X65922; CAA46715.1; ALT_INIT.
DR EMBL; D26610; BAA05655.1; ALT_INIT.
DR EMBL; D33715; AAA81564.1; ALT_INIT.
DR EMBL; U41499; AAA83776.1; ALT_INIT.
DR EMBL; U72543; AAB52915.1; ALT_INIT.
DR PIR; JCI278; JCI278.
DR GeneW; HGNC:3597; FAU.
DR MIM; 134690; -.
DR MGD; MGI:102547; Fau.
DR RIBOSOMAL PROTEIN.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 59 AA; 6648 MW; 012AC1FB55B01A4 CRC64;
Query Match 100.0%; Score 59; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.2e-53;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KHGSLARACKVRGQPPKVAKKKKKTKGRARRRMQYNRRFVNVPTFGKKKGNPANS 59
DB 1 KHGSLARACKVRGQPPKVAKKKKKTKGRARRRMQYNRRFVNVPTFGKKKGNPANS 59
RESULT 2
RS30.ARAATH STANDARD; PRT; 62 AA.
AC P49689; O82203; G9M0E4;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 40S ribosomal protein S30.
GN (RPS30A OR AT2G19750 OR F6F22.22) AND
GN (RPS30B OR AT4G29390 OR F17A13.210) AND
GN (RPS30C OR AT5G56670 OR M1K19.12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

RA Lebas M., Regad F., Iescur B.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (RPS30A).
RC STRAIN=cv. Columbia;
RX MEDLINE-20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carreira A.J., Cressy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana";
RL Nature 402:761-768(1999).
RN [3]
RP SEQUENCE FROM N.A. (RPS30B).
RC STRAIN=cv. Columbia;
RX MEDLINE-20083488; PubMed=10617198;
RA Mayer K.F.X., Scheller C., Mambut R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Enlian K.-D., Terry N.,
RA Harris B., Ansoore W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Maché R., Mueller M.,
RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohensei J., Zimmermann W., Weiler H., Rüdley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA van der Schuren J., Glynoprez B., Chung Y.-J., Vandenbusche F.,
RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,
RA Weizenecker T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Wooljman P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernerster S., Hempel S., Feldpausch M., Lambart R., Van den Daele H.,
RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,
RA van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., Mcclay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderrath K., Dauner D., Herzl A.,
RA Neumann O., Argiridou A., Vitale D., Liguori R., Plavandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schütz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton L., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Splich J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Droe K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Watere A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
RA Grant S., Shohdy N., Hasegawa A., Hamed A., Lohdl M., Johnson A.,
RA Chen E., Maria M., Martienssen R., McCombe W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana";
RL Nature 402:769-777(1999).
RN [4]
RP SEQUENCE FROM N.A. (RPS30C).
RC STRAIN=cv. Columbia;
RX MEDLINE-98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;

```

RT      *Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT      Sequence features of the regions of 1,367,185 bp covered by 19
RT      physically assigned P1 and TAC clones.
RT      DNA Res. 5:203-216(1998).
RN      [5]
RN      SEQUENCE FROM N.A.
RC      SHINOZAKI K., DAVIS R.W., ECKER J.R., THEOLOGIS A.;
RA      SHINOZAKI K., DAVIS R.W., ECKER J.R., THEOLOGIS A.;
RT      "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT      SSP consortium (Salk/Stanford/PGEC).";
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Z26869; CAA81482.1; -
DR      EMBL: AC005169; AAC62141.2; -
DR      EMBL: AL161574; CAB79697.1; -
DR      EMBL: AL096692; -; NOT_ANNOTATED_CDS.
DR      EMBL: AB013392; BAB09885.1; -
DR      EMBL: AY052341; AAK6533.1; -
DR      EMBL: AY061910; AAL31237.1; -
KW      Ribosomal protein; Multigene family.
FT      CONFLICT 17 19 QTP -> RHQ (IN REF. 1).
SQ      SEQUENCE 62 AA; 6887 MW; 95D8F3EB72F53F33 CRC64;

Query Match
Best Local Similarity 37.3%; Score 22; DB 1; Length 62;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVHGSILRAGKVRGQTPKVAQK 22
DB      3 KVHGSILRAGKVRGQTPKVAQK 24

RESULT 3
RS30_ORYLA STANDARD; PRT; 59 AA.
AC      Q9W6Y0;
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      40S ribosomal protein S30.
GN      FAU OR RPS30.
OS      Oryzias latipes (Medaka fish) (Japanese ricefish).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC      Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX      NCBI_TaxID=8090;
RN      [1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=20070552; PubMed=10603084;
RA      Henrich T., Witbrodt J.;
RT      "An in situ hybridization screen for the rapid isolation of
RT      differentially expressed genes.";
RL      Dev. Genes Evol. 210:28-33(2000).
CC      -1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
CC      C-TERMINAL EXTENSION PROTEIN (CEP) OF A UBQUITIN-LIKE PROTEIN.
CC      -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AJ238274; CAB40969.1; ALT_INT.
KW      Ribosomal protein.
SQ      SEQUENCE 59 AA; 6660 MW; E91B9C088C7772CF CRC64;

Query Match
Best Local Similarity 30.5%; Score 18; DB 1; Length 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      37 QYNRRFVNVPTEGKKKG 54
DB      37 QYNRRFVNVPTEGKKKG 54

RESULT 4
RS30_PLAFA STANDARD; PRT; 58 AA.
AC      Q96269;
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      40S ribosomal protein S30.
GN      RPS30 OR PFB0885W.
OS      Plasmodium falciparum.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5833;
RN      [1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=99021743; PubMed=9804551;
RA      Gardner M.J., Tetteil H., Carucci D.J., Cummings L.M., Aravind L.,
RA      Koehn E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
RA      Shen K., Jing J., Aston C., Lal Z., Schwartz D.C., Percec M.,
RA      Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA      Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT      "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT      falciparum.";
RL      Science 282:1126-1132(1998).
CC      -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AE001422; AAC71966.1; -
KW      Ribosomal protein.
SQ      SEQUENCE 58 AA; 6488 MW; 525675AE74D04E5F CRC64;

Query Match
Best Local Similarity 20.3%; Score 12; DB 1; Length 58;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVHGSILRAGKV 12
DB      3 KVHGSILRAGKV 14

RESULT 5
RS30_SCHPO STANDARD; PRT; 61 AA.
AC      Q42952; O14314;
DT      15-DEC-1998 (Rel. 37, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      40S ribosomal protein S30.
GN      (RPS30A OR RPS30 OR SPAC19B12.04) OR (RPS30B OR SPBC1967.03C).
OS      Schizosaccharomyces pombe (fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;

```

```

OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Lemaers G., Perret E., Bonnet C., Caput D., Picard A.;
RT "transcription of ribosomal genes is down regulated by ammonium
starvation in fission yeast.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (RPS30A AND RPS30B).
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gellies S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Shelton J., Slimmons M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicikert G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gebel C., Fuchs I., Filtz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Beer F., Zimmermann W., Medler H., Mambult R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shipkovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR S30 IN S. POMBE.
CC -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: AJ002731; CAA05693.1; -
DR EMBL: AL390814; CAC00552.1; -
DR EMBL: AL021839; CAA17057.2; -
KW Ribosomal protein; Multigene family
SQ SEQUENCE 61 AA; 6910 MW; 4B9C17AFLA326F22 CRC64;

Query Match 20.3%; Score 12; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKV 12
DB 3 KVGSLARAGKV 14

```

```

GN (RPS30A OR YLR287BC OR L8003.23) AND (RPS30B OR YOR182C).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE (RPS30A).
RC STRAIN=YRB141;
RX MEDLINE=96278780; PubMed=8662789;
RA Baker R.T., Williamson N.A., Wettenhall R.E.H.;
RT "The yeast homolog of mammalian ribosomal protein S30 is expressed
from a duplicated gene without a ubiquitin-like protein fusion
sequence. Evolutionary implications.";
RL J. Biol. Chem. 271:13549-13555(1996).
RN [2]
RP SEQUENCE FROM N.A. (RPS30B).
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Tatch A., Trevaakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (RPS30B).
RA Hughes B., Pohl T.M.;
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- MASS SPECTROMETRY: MW=6987; MW_ERR=3.4; METHOD=MALDI.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR S30 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: U48700; AAC49317.1; -
DR EMBL: U48699; AAC49316.1; -
DR EMBL: U83406; AAB41050.1; -
DR EMBL: U83407; AAB41051.1; -
DR EMBL: U17243; AAB67333.1; -
DR EMBL: Z75090; CAA9391.1; -
DR SGD: S0004278; RPS30A.
DR SGD: S0005708; RPS30B.
KW Ribosomal protein; Multigene family.
KW INT_MET 0
FT SEQUENCE 62 AA; 6987 MW; 65EADAE8D9FC586 CRC64;

Query Match 20.3%; Score 12; DB 1; Length 62;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVGSLARAGKV 12
DB 2 KVGSLARAGKV 13

```

```

RESULT 7
RS30_ONCMY
ID RS30_ONCMY STANDARD; PRT; 11 AA.
AC P83328;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 40S ribosomal protein S30 (Fragment).
GN FNU.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

```

```

OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Skin mucus;
RA Fernandes J.M.O., Smith V.J.;
RT "A novel anticribobial function for a ribosomal peptide from skin
  secretions of rainbow trout.";
RL Submitted (Apr-2002) to the SWISS-PROT data bank.
CC -1- FUNCTION: Has antibacterial activity against Gram-positive
  bacteria.
CC -1- MASS SPECTROMETRY: MW=6676.6; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
KW Ribosomal protein; Antibiotic.
FT NON_TER
SQ SEQUENCE 11 AA; 1123 MW; 2312AB630DD735B8 CRC64;

Query Match 18.6%; Score 11; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYHGSILARAGK 11
DB 1 KYHGSILARAGK 11

RESULT 8
HNS_PROVU
ID HNS_PROVU STANDARD; PRT; 133 AA.
AC P18818;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE DNA-binding protein H-NS (histone-like protein HLP-II).
GN HNS OR HNSA.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxId=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89171270; PubMed=2494066;
  la Teana A., Falconi M., Scarlato V., Iammi M., Pon C.L.;
  "Characterization of the structural genes for the DNA-binding protein
  H-NS in Enterobacteriaceae.";
  FEBS Lett. 244:34-38(1989).
CC -1- FUNCTION: H-NS BINDS TIGHTLY TO DS-DNA, INCREASES ITS THERMAL
  STABILITY AND INHIBITS TRANSCRIPTION. IT ALSO BINDS TO SS-DNA AND
  RNA BUT WITH A MUCH LOWER AFFINITY. H-NS HAS POSSIBLE HISTONE-LIKE
  FUNCTION. MAY BE A GLOBAL TRANSCRIPTIONAL REGULATOR THROUGH ITS
  ABILITY TO BIND TO CURVED DNA SEQUENCES, WHICH ARE FOUND IN
  REGIONS UPSTREAM OF A CERTAIN SUBSET OF PROMOTERS. IT PLAYS A ROLE
  IN THE THERMAL CONTROL OF PILI PRODUCTION. IT IS SUBJECT TO
  TRANSCRIPTIONAL AUTO-REPRESSION. IT BINDS PREFERENTIALLY TO THE
  UPSTREAM REGION OF ITS OWN GENE RECOGNIZING TWO SEGMENTS OF DNA ON
  BOTH SIDES OF A BEND CENTERED AROUND -150 (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE HISTONE-LIKE PROTEIN H-NS FAMILY.
DR HSP; P08936; 1HNH.
DR InterPro; IPR001801; Histone_HNS.
DR Pfam; PF00816; Histone_HNS; 1.
DR ProDom; PD007337; Histone_HNS; 1.
DR SMART; SM00528; HNS; 1.
KM DNA-binding; transcription regulation; Repressor.
FT INIT_MET
SQ SEQUENCE 133 AA; 15134 MW; B92C49F7995C5ED5 CRC64;

Query Match 11.9%; Score 7; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 KTGRRAKR 34

```

```

DB 82 KTGRRAKR 88

RESULT 9
YOC2_CAEL
ID YOC2_CAEL STANDARD; PRT; 346 AA.
AC 009261;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 39.2 kDa protein C32A3.2 in chromosome III.
GN C32A3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Felodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Thomas K.;
RL Submitted (Feb-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
  -----
DR EMBL: 248241; CAA88285.1; -
DR WormPeP: C32A3.2; CE01506.
KW Hypothetical protein.
SQ SEQUENCE 346 AA; 39221 MW; D511225F5062B420 CRC64;

Query Match 11.9%; Score 7; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 EKKKKKT 29
DB 6 EKKKKKT 12

RESULT 10
COLB_HUMAN
ID COLB_HUMAN STANDARD; PRT; 489 AA.
AC Q9BR76;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coronin 1B (Coronin 2).
GN CORO1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strusberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN CYTOKINESIS, MOTILITY, AND SIGNAL
  TRANSDUCTION (BY SIMILARITY).
CC -1- SUBUNIT: BINDS ACTIN (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE CORONIN FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: BC006449; AAA06449.1; -.

DR Genew; HGNC:2253; CORO1B.

DR InterPro: IPR001680; WD40.

DR Pfam: PF004400; WD40; 3.

DR Prodom: PD000018; WD40; 1.

DR SMART: SM00320; WD40; 3.

DR PROSITE: PS00678; WD_REPEATS_1; 1.

DR PROSITE: PS0082; WD_REPEATS_2; 2.

DR PROSITE: PS0294; WD_REPEATS_REGION; 1.

DR Actin-binding; Repeat; WD repeat; Coiled coil.

FT REPEAT 80 120 WD 1.

FT REPEAT 130 170 WD 2.

FT REPEAT 174 213 WD 3.

FT REPEAT 217 260 WD 4.

FT REPEAT 265 305 WD 5.

FT DOMAIN 449 474 COILED COIL (POTENTIAL).

SQ SEQUENCE 489 AA; 54234 MW; A6012FDA683EC859 CRC64;

Query Match 11.9%; Score 7; DB 1; Length 489;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GSILRAG 10
|||||||

DB 442 GSILRAG 448

RESULT 11
POLY_EEYVM PRT: 1254 AA.
ID P36331; Q66587; Q66588; Q66590; Q66591;
AC P36331-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 40, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (P130) [contains: Coat protein C (EC 3.4.21.-)
DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS Venezuelan equine encephalitis virus (strain Mena II).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=36384;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93187617; PubMed-8445371;
RA Snelder J.M., Kinney R.M., Tsuchiya K.R., Trent D.W.;
RT "Molecular evidence that epizootic Venezuelan equine encephalitis
RT (VEE) I-AB viruses are not evolutionary derivatives of enzootic VEE
RT subtype I-E or II viruses."
RL J. Gen. Virol. 74:519-523(1993).
CC -1- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THE 6 KDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: L04599; AAA42990.1; -.

DR EMBL; L04599; AAA42989.1; ALT_TERM.

DR EMBL; L04599; AAA42991.1; ALT_SEQ.

DR EMBL; L04599; AAA42992.1; ALT_SEQ.

EMBL; L04599; AAA42993.1; ALT_SEQ.

DR EMBL; L04599; AAA42994.1; ALT_INIT.

DR PIR; J01979; J01979.

DR HSSP; P03315; IYCP.

DR MEROPS; S03.001; -.

DR InterPro: IPR002548; Alpha_E1_glycop.

DR InterPro: IPR000936; Alpha_E2_glycop.

DR InterPro: IPR002533; Alpha_E3_glycop.

DR InterPro: IPR001836; Alpha_core.

DR InterPro: IPR000930; Togavirin.

DR Pfam: PF00943; Alpha_E2_glycop; 1.

DR Pfam: PF00944; Alpha_core; 1.

DR Pfam: PF01563; Alpha_E3_glycop; 1.

DR Pfam: PF01589; Alpha_E1_glycop; 1.

DR PRINTS; PR00798; TOGAVIRIN.

DR Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;

KW Serine protease.

FT CHAIN 1 274 COAT PROTEIN C.

FT CHAIN 275 333 SPIKE GLYCOPROTEIN E3.

FT CHAIN 334 756 SPIKE GLYCOPROTEIN E2.

FT CHAIN 757 812 6 KDA PEPTIDE.

FT CHAIN 813 1254 SPIKE GLYCOPROTEIN E1.

FT ACT_SITE 157 157 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 157 157 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 225 225 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT TRANSSEM 701 721 POTENTIAL.

FT TRANSSEM 794 813 POTENTIAL.

FT TRANSSEM 1231 1248 POTENTIAL.

FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 1254 AA; 138343 MW; FB9DE8265F20211 CRC64;

Query Match 11.9%; Score 7; DB 1; Length 1254;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 KKKKKTG 30
|||||||

DB 88 KKKKKTG 94

RESULT 12
R27A_DICDI STANDARD; PRT: 78 AA.
ID R27A_DICDI
AC P14757;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S ribosomal protein s27a.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44869;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89352609; PubMed-2548604;
RA Omachi T., Giorda R., Shaw D.R., Ennis H.L.;
RT "Molecular organization of developmentally regulated Dictyostelium
RT discoideum ubiquitin CDNA's."
RL Biochemistry 28:5226-5231(1989).
CC -1- MISCELLANEOUS: THIS RIBOSOMAL PROTEIN IS SYNTHESIZED AS A
CC C-TERMINAL EXTENSION PROTEIN (CEP) OF UBLOUTIN.
CC -1- SIMILARITY: BELONGS TO THE S27AE FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M23750; AAA33264.1; ALT_INIT.
 DR PIR: E34080; UODOR7.
 DR DictyDb: DD01026; .
 DR InterPro: IPR002906; Ribosomal_S27.
 DR Pfam: PF01599; Ribosomal_S27; 1.
 KW Ribosomal protein; Zinc-finger; Metal-binding.
 FT DOMAIN 1 24 Lys-RICH (HIGHLY BASIC).
 FT ZN-FING 46 69 C4-TYPE.
 SQ SEQUENCE 78 AA: 8915 MW: F77B3E09D1B7246F CRC64;
 Query Match 10.2%; Score 6; DB 1; Length 78;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 24 KKKKK 29
 Db 4 KKKKK 9
 RESULT 13
 RS18_BACSU STANDARD; PRT; 78 AA.
 AC P21475;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S18 (BS21).
 GN RPSR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96051385; PubMed=7584024;
 RA Ogasawara N., Nakai S., Yoshikawa H.;
 RT "Systematic sequencing of the 180 kilobase region of the Bacillus
 RL DNA Res. 1:1-14(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaser P., Goffeau A., Goldightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halsech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maucel C., Medigue C.,
 RA Medina N., Meliade R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Pressan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
 RA Setoichi J., Sekowska A., Seror S.J., Serro P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassart A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Wetzenecker T.,
 RA Winters P., Wipit A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa K., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";

RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE OF 1-23.
 RX MEDLINE=82219212; PubMed=6806564;
 RA Higo K.I., Otake E., Osawa S.;
 RT "Purification and characterization of 30S ribosomal proteins from
 RL Bacillus subtilis: correlation to Escherichia coli 30S proteins.";
 CC Mol. Gen. Genet. 185:239-244(1982).
 CC -1- FUNCTION: This protein has been implicated in aminocycl-transfer
 CC RNA binding. It appears to be situated at the decoding site of
 CC messenger RNA.
 CC -1- SIMILARITY: BELONGS TO THE S18P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: D26185; BA05219.1; ALT_INIT.
 DR EMBL: Z99124; CAB16126.1; ALT_INIT.
 DR PIR: S11368; S11368.
 DR Subtilist; BG10047; rpsr.
 DR InterPro: IPR001648; Ribosomal_S18.
 DR Pfam: PF01084; Ribosomal_S18; 1.
 DR PRINTS: PR00974; RIBOSOMALS18.
 DR ProDom: PD002239; Ribosomal_S18; 1.
 DR TIGRFAMs: TIGR00165; S18; 1.
 DR PROSITE: PS00057; RIBOSOMAL_S18; 1.
 KW Ribosomal protein; RNA-binding; Complete proteome.
 FT INT_MEF 0 0
 SQ SEQUENCE 78 AA: 8838 MW: 1CF19CE039C64FF2 CRC64;
 Query Match 10.2%; Score 6; DB 1; Length 78;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 30 GAKRR 35
 Db 7 GAKRR 12
 RESULT 14
 EX7S_VIBCH STANDARD; PRT; 80 AA.
 ID EX7S_VIBCH
 AC O9KTL1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
 DE (Exonuclease VII small subunit).
 GN XSEB OR VC0891.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L.V., Uitterlinden T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE
 CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER

|||||
Db 11 SLARAG 16

Search completed: January 9, 2003, 12:29:50
Job time : 11 secs

CC INFO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonuclease cleavage in either 5' to 3' or 3' to 5' direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE004173; AAF94053.1; -.
CC TIGR: VC0891; -.
CC DR InterPro: IPR003761; Exonuc_VII_S.
CC DR Pfam: PF02609; Exonuc_VII_S; 1.
CC DR TIGRFAMs: TIGR01280; xseB; 1.
CC DR Hydrolase; Nuclease; Exonuclease; Complete proteome.
CC SEQUENCE 80 AA; 8911 MW; F9B/F5A9C6821DD1 CRC64;

Query Match 10.2%; Score 6; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 SLARAG 10
|||||
Db 44 SLARAG 49

RESULT 15
YGRM_MICEC
ID YGRM_MICEC STANDARD; PRT; 105 AA.
AC P24620;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in grm 3' region (Fragment).
OS Micromonospora echinospora (Micromonospora purpurea).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Micromonosporineae; Micromonosporaceae;
OC Micromonospora.
OX NCBI_TaxID=1877;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91192615; PubMed=2013410;
RA Kelenen G.H., Cundliffe E., Financsek I.;
RT "Cloning and characterization of gentamicin-resistance genes from Micromonospora purpurea and Micromonospora rosea".
RL Gene 98:53-60(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M55520; AAA25337.1; -.
CC DR PIR: PM0017; PM0017.
CC DR Antibiotic resistance; Hypothetical protein.
CC KM NON_TER 105
CC FT NON_TER 105
CC SEQUENCE 105 AA; 11111 MW; 0B398039381DB012 CRC64;

Query Match 10.2%; Score 6; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 SLARAG 10

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 9, 2003, 12:28:19 : Search time 29 seconds
(without alignments)
419.199 Million cell updates/sec

Title: US-09-424-815E-1

Perfect score: 59

Sequence: 1 KHGSLARAGKVRGOTPKVA.....RRFVNVVPTFGKKKGPNNNS 59

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	133	4 Q9H5V4	Q9H5V4 homo sapien
2	59	100.0	133	4 Q9J24	Q9J24 mus musculu
3	59	100.0	137	11 Q920W8	Q920W8 mus musculu
4	59	100.0	137	11 Q920W7	Q920W7 mus spicille
5	59	100.0	137	11 Q91V99	Q91V99 mus musculu
6	23	39.0	133	13 Q90YPI	Q90YPI ictalurus p
7	22	37.3	62	10 Q9W0E4	Q9W0E4 arabidopsis
8	21	35.6	230	4 Q15351	Q15351 homo sapien
9	15	25.4	132	5 Q9VDH8	Q9VDH8 drosophila
10	13	22.0	130	5 Q18231	Q18231 caenorhabdi
11	12	20.3	131	5 Q96201	Q96201 spodoptera
12	9	15.3	309	10 Q9S930	Q9S930 arabidopsis
13	8	13.6	135	16 Q971R7	Q971R7 clostridium
14	7	11.9	134	2 Q91657	Q91657 proteus mlr
15	7	11.9	154	16 Q9RS55	Q9RS55 deinococcus
16	7	11.9	157	5 Q9NB4	Q9NB4 drosophila

17	7	11.9	173	10 Q920B1	Q920B1 arabidopsis
18	7	11.9	198	2 Q9AGU8	Q9AGU8 bruceella ab
19	7	11.9	215	11 Q64075	Q64075 rattus sp.
20	7	11.9	224	10 Q9LIR5	Q9LIR5 arabidopsis
21	7	11.9	233	10 Q9F276	Q9F276 arabidopsis
22	7	11.9	233	10 Q9C9C6	Q9C9C6 arabidopsis
23	7	11.9	233	10 Q9C9C5	Q9C9C5 arabidopsis
24	7	11.9	259	16 Q9AD23	Q9AD23 streptomyces
25	7	11.9	283	4 Q9NSK3	Q9NSK3 homo sapien
26	7	11.9	309	5 Q95T41	Q95T41 drosophila
27	7	11.9	314	16 Q66535	Q66535 aquilex aeo
28	7	11.9	339	5 Q95XC7	Q95XC7 caenorhabdi
29	7	11.9	380	4 Q96GN4	Q96GN4 homo sapien
30	7	11.9	380	5 Q9W0W6	Q9W0W6 drosophila
31	7	11.9	381	4 Q15694	Q15694 homo sapien
32	7	11.9	424	16 Q9S2F0	Q9S2F0 streptomyces
33	7	11.9	429	2 Q9X4J1	Q9X4J1 streptomyces
34	7	11.9	454	16 Q9K509	Q9K509 bacillus ha
35	7	11.9	542	13 Q91290	Q91290 pleurodeles
36	7	11.9	576	4 Q9NXU4	Q9NXU4 homo sapien
37	7	11.9	594	16 Q98C19	Q98C19 rhizobium l
38	7	11.9	806	10 Q9M8T5	Q9M8T5 arabidopsis
39	7	11.9	863	10 Q9W0D8	Q9W0D8 arabidopsis
40	7	11.9	1254	12 Q9WC26	Q9WC26 venezuelan
41	7	11.9	1254	12 Q9YK03	Q9YK03 venezuelan
42	7	11.9	1254	12 Q88979	Q88979 venezuelan
43	7	11.9	1254	12 Q8V293	Q8V293 venezuelan
44	7	11.9	1254	12 Q8V291	Q8V291 venezuelan
45	7	11.9	1264	12 Q8UYH1	Q8UYH1 venezuelan

ALIGNMENTS

RESULT 1

ID	Q9H5V4	PRELIMINARY:	PRT:	133 AA.
AC	Q9H5V4:			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	CDNA: FLJ22986 fis, clone KAT11742.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RA	Matenabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,			
RA	Suzuki Y., Odayashi M., Nishii T., Shibahara T., Tanaka T.,			
RA	Nakamura Y., Isonagi T., Sugano S.;			
RT	"NEDO human cDNA sequencing project."			
RL	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AK026639; BAB15515.1; -			
DR	HSSP: P02248; IUBI.			
DR	InterPro: IPR000526; Ubiqultin.			
DR	Pfam: PF00240; ubiqultin; 1.			
DR	PRINTS: PR00348; UBIQUITIN.			
DR	SMART: SM00213; UBO; 1.			
DR	PROSITE: PS00299; UBIQUITIN_1; 1.			
DR	PROSITE: PS50053; UBIQUITIN_2; 1.			
SO	SEQUENCE 133 AA; 14390 MW; 5D2F81F2A355B559 CRC64;			

Query Match 100.0%; Score 59; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.5e-56;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 KHGSLARAGKVRGOTPKVAQEKKKKTKGRKRRQYRRRVNVVPTFGKKKGPNNNS 59
DB	75 KHGSLARAGKVRGOTPKVAQEKKKKTKGRKRRQYRRRVNVVPTFGKKKGPNNNS 133

RESULT 2

```

09J24 ID 09J24 PRELIMINARY; PRT; 133 AA.
AC 09J24:
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Monoclonal non-specific suppressor factor beta (Pindel-Biskis-Reilly
DE murine sarcoma virus (FBR-MuSV) ubiquitously expressed) (fox
DE derived).
GN FAV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS OB.
RA Nie G.-Y., Li Y., Salmonsens L.A., Clements J.A., Findlay J.K.;
RT "Identification of monoclonal non-specific suppressor factor beta as
RT one of the genes differentially expressed at implantation sites
RT compared to interimplantation sites in the mouse uterus.";
RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE, AND KIDNEY;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konno S., Yamahata I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gunclich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monhaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AF147745; AAF80246.1; -
DR EMBL: AK008466; BAB25684.1; -
DR EMBL: AK002355; BAB2034.1; -
DR HSSP: P02248; IUBI.
DR MGD: MGI:102547; Fau.
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00240; ubiquitin; 1.
DR PRINTS: PRO0346; UBIOUITIN.
DR SMART: SM00213; UBO; 1.
DR PROSITE: PS00299; UBIOUITIN_1; 1.
DR PROSITE: PS50053; UBIOUITIN_2; 1.
SQ SEQUENCE 133 AA; 14416 MW; 20B7C774A022AB16 CRC64;

Query Match 100.0%; Score 59; DB 11; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.5e-56;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

09Y 1 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 59
DB 75 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 133

RESULT 3
0920W8 ID 0920W8 PRELIMINARY; PRT; 137 AA.
AC 0920W8:
DT 01-DEC-2001 (TReMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Fau protein (Fragment).
GN FAV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BLG2/MSF;
RA Liu Y., Kitano T., Koide T., Shirolshi T., Moriwaiki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039086; BAB68610.1; -
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00240; ubiquitin; 1.
DR PROSITE: PS00299; UBIOUITIN_1; UNKNOWN_1.
DR PROSITE: PS50053; UBIOUITIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 137 AA; 14845 MW; 57099FF4055D8B2B CRC64;

Query Match 100.0%; Score 59; DB 11; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.6e-56;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

09Y 1 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 59
DB 79 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 137

RESULT 4
0920W7 ID 0920W7 PRELIMINARY; PRT; 137 AA.
AC 0920W7:
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Fau protein (Fragment).
GN FAV.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZBN;
RA Liu Y., Kitano T., Koide T., Shirolshi T., Moriwaiki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039093; BAB68617.1; -
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00240; ubiquitin; 1.
DR PROSITE: PS00299; UBIOUITIN_1; UNKNOWN_1.
DR PROSITE: PS50053; UBIOUITIN_2; 1.
FT NON_TER 1
SQ SEQUENCE 137 AA; 14787 MW; 57099FF7065D8828 CRC64;

Query Match 100.0%; Score 59; DB 11; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.6e-56;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

09Y 1 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 59
DB 79 KVHGSILARACKVVGQTPKVAKOEKKKKKTGRAKRRMOYNRRFVNVVPTFGKKKGPNNNS 137

RESULT 5
091V99 ID 091V99 PRELIMINARY; PRT; 137 AA.
AC 091V99:

```

DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Fau Protein (Fragment).
GN FAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS;
RA Liu Y., Kitano T., Shiroishi T., Moriwaki K., Saitou N.;
RT "Conspicuous differences among Gene Genealogies of 21 Nuclear Genes of
RT Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB039084; BAB68608.1; -
DR EMBL; AB039085; BAB68609.1; -
DR EMBL; AB039087; BAB68611.1; -
DR EMBL; AB039088; BAB68612.1; -
DR EMBL; AB039089; BAB68613.1; -
DR EMBL; AB039090; BAB68614.1; -
DR EMBL; AB039091; BAB68615.1; -
DR EMBL; AB039092; BAB68616.1; -
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00240; Ubiquitin; 1.
DR PROSITE: PS00289; UBIQUITIN_1; UNKNOWN_1.
DR PROSITE: PS50053; UBIQUITIN_2; 1.
FT NON_TER
FT 1
SQ SEQUENCE 137 AA; 14787 MW; 57099FF7065D8828 CRC64;

Query Match 100.0%; Score 59; DB 11; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.6e-56;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVHGSILARAGKVRGOTPKVAKOEEEEKKTGRKRRMQYNNRFVNVPTFGKKKGNPANS 59
DB 79 KVHGSILARAGKVRGOTPKVAKOEEEEKKTGRKRRMQYNNRFVNVPTFGKKKGNPANS 137

RESULT 6
Q90YPI PRELIMINARY; PRT; 133 AA.
AC Q90YPI;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE 40S ribosomal protein S30.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Karsl A., Patterson A., Feng J., Liu Z.J.;
RT "Translational machinery of channel catfish: I. A transcriptomic
RT approach to the analysis of 32 40S ribosomal protein genes and their
RT expression.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402841; AAK95215.1; -
DR InterPro: IPR000626; Ubiquitin.
DR Pfam: PF00240; Ubiquitin; 1.
DR PROSITE: PS50053; UBIQUITIN_2; 1.
DR Ribosomal protein.
KW Ribosomal protein.
SQ SEQUENCE 133 AA; 14504 MW; 62036BB0E72C5C64; CRC64;

Query Match 39.0%; Score 23; DB 13; Length 133;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 QYNRRFVNVPTFGKKKGNPANS 59
|||||

DB 111 QYNRRFVNVPTFGKKKGNPANS 133

RESULT 7
Q9M0E4 PRELIMINARY; PRT; 62 AA.
AC Q9M0E4;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Ribosomal protein S30 homolog (40S ribosomal protein S30 homolog)
DE (A02919750/F6F22.22).
GN ATG29390.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA.
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:203-216(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Saitou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Saitou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161574; CAB79697.1; -
DR EMBL; AB013392; BAB09885.1; -
DR EMBL; AY052341; AAK96533.1; -
DR EMBL; AY061910; AAL31237.1; -
DR Ribosomal protein.
KW Ribosomal protein.
SQ SEQUENCE 62 AA; 6887 MW; 95D8F3EB72F53F33 CRC64;

Query Match 37.3%; Score 22; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.4e-16;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVHGSILARAGKVRGOTPKVAKO 22
|||||

```

Db      3  KHGSLRACKRGQOTPKVAKO 24
      |||||||
RESULT 8
Q13531 ID Q13531 PRELIMINARY; PRT; 230 AA.
AC Q13531;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE SBR4B (Fragment).
GN SBR4B(HUMAN).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RA Ruehlmann A., Gupta A., Terhorst C.;
RT "A novel murine RRM-type protein and its human homolog ";
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; X75315; CAAS3064.1; -.
DR HSSP; P09651; 10P1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
DR NON_TER 1 1
SQ SEQUENCE 230 AA; 25220 MW; C747D6500608461 CRC64;

Query Match 35.6%; Score 21; DB 4; Length 230;
Best Local Similarity 100.0%; Pred. No. 1e-14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 RRMQYNNRRFVNVPTFGKKKG 54
Db 1 RRMQYNNRRFVNVPTFGKKKG 21

RESULT 9
Q9VDH8 ID Q9VDH8 PRELIMINARY; PRT; 132 AA.
AC Q9VDH8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE CG15697 protein (RH08962p).
GN CG15697.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ayrill J.F., Agbayani A., An H.-U., Andrews-Plannkoch C., Baldwin D.,
RA Bielew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Krawitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacled J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weissstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Fartan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacled J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003732; AAF55815.1; -.
DR Flybase; FBgn0038834; CG15697.
DR InterPro; IPR000626; Ubligutlin.
DR Pfam; PF00240; ublignitin; 1.
SQ SEQUENCE 132 AA; 14585 MW; E07C754F3316FDE5 CRC64;

Query Match 25.4%; Score 15; DB 5; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 KOEKKKKKTGRARR 35
Db 93 KOEKKKKKTGRARR 107

RESULT 10
Q18231 ID Q18231 PRELIMINARY; PRT; 130 AA.
AC Q18231;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C26F1.4 protein.
GN C26F1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsouh R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier W., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

```

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smadon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weisscock L., Wilkinson-Sproat J., Woldman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Geisel C., Steliyes L., Bradshaw H.;
RT "The sequence of C. elegans cosmid C26F1.";
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (Nov-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U53148; AAB37076.1; -.
DR HSSP: P02248; IUBI.
DR InterPro: IPR000526; Ubiquitin.
DR Pfam: PF00240; Ubiquitin; 1.
DR SMART: SM00213; UBO; 1.
SQ SEQUENCE 130 AA; 14033 MW; 50DC09AFB9F48532 CRC64;
Query Match 22.0%; Score 13; DB 5; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KYHGSLARAGKV 13
DB 72 KYHGSLARAGKV 84

RESULT 11
Q96201 PRELIMINARY; PRT; 131 AA.
ID 096201;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ribosomal protein S30.
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuidae; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RN RP SEQUENCE FROM N.A.
RA Landais I., Ogliastro M., Mita K., Nohata J., Lopez-Ferber M.,
RA Duonot-Cerutti M., Fournier P., Devauchelle G.;
RT "Full-length ribosomal protein sequence from an EST library of
RT Spodoptera frugiperda cells (Sf9).";
RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF400225; AAK92197.1; -.
SQ SEQUENCE 111 AA; 14314 MW; 03AEAD031EBC1B04 CRC64;
Query Match 20.3%; Score 12; DB 5; Length 131;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KYHGSLARAGKV 12
DB 73 KYHGSLARAGKV 84

RESULT 12
Q9S9J0 PRELIMINARY; PRT; 309 AA.
ID 09S9J0;
AC 09S9J0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE T23K8.1 (Fragment).
GN T23K8.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Li J., Kremenetskaia I., Ngan I., Luros J., Gonzalez A., Altrefi H.,
RA Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
RA Huizar L., Kim C., Palm C.J., Rowley D., Shinn P., Walker M.,
RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC T23K8 sequence.";
RL Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Theologis A.;
RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC007230; AAD26867.1; -.
FT NON_TER 309 309
SQ SEQUENCE 309 AA; 34049 MW; 4FF1D7AF877A0D1 CRC64;
Query Match 15.3%; Score 9; DB 10; Length 309;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 KKKKKTGRA 32
DB 232 KKKKKTGRA 240

RESULT 13
Q971R7 PRELIMINARY; PRT; 135 AA.
ID 0971R7;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Uncharacterized protein, yJDF B. subtilis ortholog.
GN CAC1573.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE-2159325; PubMed-11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng O.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,
RA Tatusov R.L., Sabatini F., Doucette-Stamm L., Soucaille P., Daly M.D.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007667; AAK79540.1; -.
KW Complete proteome.
SQ SEQUENCE 135 AA; 16388 MW; 387D5F8D11444E7A CRC64;
Query Match 13.6%; Score 8; DB 16; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 KKKKKKK 28
DB 124 KKKKKKK 131

RESULT 14

Q9L657 PRELIMINARY; PRT; 134 AA.
AC Q9L657;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Histone-like nucleoid structuring protein.
GN HNS.
OS Proteus mirabilis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HI4320;
RA Coker C., Bakare O.O., Mobley H.L.T.;
RT "H-NS is a Repressor of the Proteus mirabilis Urease Transcriptional
Activator Gene uer.";
RL J. Bacteriol. 0:0-0(2000).
DR EMBL: AF240693; AAF61240.1; -.
DR HSSP: P08936; 1HNR.
DR InterPro: IPR001801; Histone_HNS.
DR Pfam: PF00816; Histone_HNS; 1.
DR ProDom: PD007337; Histone_HNS; 1.
DR SMART: SM00528; HNS; 1.
SQ SEQUENCE 134 AA; 15249 MW; 202C49BCA12597B3 CRC64;

Query Match

Best Local Similarity 11.9%; Score 7; DB 2; Length 134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 KTGRKR 34
|||||
DB 83 KTGRKR 89

RESULT 15

Q9RSE5 PRELIMINARY; PRT; 154 AA.
AC Q9RSE5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein DR2179.
GN DR2179.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofatt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE002051; AAF11732.1; -.
DR TIGR: DR2179; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 154 AA; 16181 MW; 58AA86884AA0D4C4 CRC64;

Query Match

Best Local Similarity 11.9%; Score 7; DB 16; Length 154;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 RAGKVRG 14
|||||
DB 39 RAGKVRG 45

Search completed: January 9, 2003, 12:30:26
Job time : 30 secs